BMC Medical Informatics and Decision Making

RESEARCH

Open Access

Predictive value of machine learning for the progression of gestational diabetes mellitus to type 2 diabetes: a systematic review and meta-analysis

Meng Zhao[1], Zhixin Yao[1], Yan Zhang[1], Lidan Ma[1], Wenquan Pang[1], Shuyin Ma[2], Yijun Xu[1*] and Lili Wei[3*]

Abstract

Background This systematic review aims to explore the early predictive value of machine learning (ML) models for the progression of gestational diabetes mellitus (GDM) to type 2 diabetes mellitus (T2DM).

Methods A comprehensive and systematic search was conducted in Pubmed, Cochrane, Embase, and Web of Science up to July 02, 2024. The quality of the studies included was assessed. The risk of bias was assessed through the prediction model risk of bias assessment tool and a graph was drawn accordingly. The meta-analysis was performed using Stata15.0.

Results A total of 13 studies were included in the present review, involving 11,320 GDM patients and 22 ML models. The meta-analysis for ML models showed a pooled C-statistic of 0.82 (95% CI: 0.79 ~ 0.86), a pooled sensitivity of 0.76 (0.72 ~ 0.80), and a pooled specificity of 0.57 (0.50 ~ 0.65).

Conclusion ML has favorable diagnostic accuracy for the progression of GDM to T2DM. This provides evidence for the development of predictive tools with broader applicability.

Keywords Clinical prediction model, Gestational diabetes mellitus, Machine learning, Type 2 diabetes mellitus, Systematic review

Background

Gestational diabetes mellitus (GDM) is characterized by impaired glucose tolerance during pregnancy [1]. The incidence of GDM has increased to 20% due to improved quality of life [2]. It can induce various complications, which bring a heavy burden to GDM patients and the medical system [3]. Compared with their peers, women with a history of GDM have a higher risk of developing type 2 diabetes mellitus (T2DM) [4] and a 7-fold increased risk of T2DM later in life [5], despite having similar baseline levels of impaired glucose tolerance [6]. The incidence of T2DM can be as high as 30–50% within 5–10 years after pregnancy [7, 8]. T2DM may lead to

*Correspondence:
Yijun Xu
xuyijun@qdu.edu.cn
Lili Wei
13573828157@163.com
[1]Department of Endocrinology and Metabolic Diseases, The Affiliated Hospital of Medical College Qingdao University, Qingdao, Shandong 266003, China
[2]Department of Emergency Pediatric, The Affiliated Hospital of Medical College Qingdao University, Qingdao, Shandong 266003, China
[3]Department of Nursing, The Affiliated Hospital of Medical College Qingdao University, Qingdao, Shandong 266003, China

various complications, such as retinopathy, nephropathy, neuropathy, and vascular diseases [9, 10], and patients with a history of GDM are more likely to develop health-threatening diseases in subsequent disability-adjusted life years. The risk of T2DM can be reduced through early identification and healthy living [11–13]. A balanced diet that includes meals and snacks, fiber-rich plant foods (legumes, nuts, seeds, herbs), non-starchy vegetables, nutrient-rich carbohydrate foods (whole grains), fish, and unsaturated vegetable oils can greatly benefit a healthy life. Additionally, aerobic exercise and resistance training both improve skeletal muscle, adipose tissue, liver, and pancreatic function; self-monitoring of blood glucose and adherence to medication [14–16] are essential for patients with a history of GDM. In recent years, machine learning (ML) has been gradually applied to clinical practice due to its desirable prediction and stratification performance [17, 18], such as the diagnosis of disease status (e.g., tumor lymph node metastasis status [19, 20], and molecular subtype [21]), prognostic prediction (e.g., early tumor recurrence [22]), and prediction of treatment responses [23, 24]. The modeling techniques and variables of ML are diverse, which is one of the reasons why its predictive performance is highly controversial. Hence, this meta-analysis was conducted to explore the accuracy of ML in predicting T2DM risk in patients with a history of GDM and provide references for the development and update of simple risk stratification tools in this field.

Materials and methods

This paper followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statements [25] and was registered in the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42022333398).

Search strategy

Pubmed, Cochrane, Embase, and Web of Science were comprehensively searched for relevant studies until May 25, 2023, without restrictions on languages and regions. The combination of subject headings and free words was used as the strategy. An additional search of the databases was conducted on July 02, 2024 to minimize the risk of missing new publications. The specific search process is shown in the supplementary material 1.

Inclusion criteria

The inclusion criteria were as follows: (i) study subjects were women with GDM; (ii) studies explicitly described whether GDM occurred during the follow-up period; (iii) in studies of ML, any of the following metrics (Roc, C-statistic, sensitivity/recall, specificity, accuracy, precision, confusion matrix, diagnostic four-compartment table, F1 score) were used to assess the accuracy of ML models. Therefore, at least one of these metrics was included for evaluating the accuracy of ML models in predicting T2DM in GDM patients.

Exclusion criteria

Meanwhile, the studies were excluded for the following reasons: (i) systematic reviews or meta-analyses, consensus statements, guidelines, review articles, and conference abstracts published without peer review; (ii) only risk factors were analyzed without building a complete prediction model; (iii) non-English articles.

Literature screening and data extraction

All retrieved literature was imported into EndNote X9. After duplicates were removed, the titles and abstracts of the remaining studies were preliminarily reviewed to exclude ineligible literature before the full texts were downloaded. After a full-text review, 13 studies were finally included.

A standard spreadsheet was prepared for information extraction, including first author, publication year, country, study type, sample source, sample size of T2DM in the training set, sample size in the training set, external validation, sample size of T2DM in the validation set, sample size in the validation set, method for variable selection, model type, number of model variables, model variables, outcomes, and follow-up duration.

Literature screening and information extraction were performed independently by two researchers (Zhao M and Wei LL) and the results were subsequently cross-checked. Any disputes were tackled through discussion with a third researcher (Ma LD).

Quality assessment

The risk of bias (ROB) of ML models in eligible studies was evaluated using the Prediction Model Risk of Bias Assessment Tool (PROBAST) in four major domains: participants, predictors, outcomes, and statistical analysis. The four domains included two, three, six, and nine questions, respectively. These questions were answered as yes/probable yes (Y/PY), no/probably no (N/PN), or no information (NI). If a domain was answered with at least one N/PN, it was rated as having high risk. In contrast, a domain with all its questions answered as Y/PY was regarded as having low risk. When all domains were at low risk, the overall ROB was rated as low. However, the overall ROB was high when at least one domain was considered at high risk.

Statistical analysis

Data synthesis

A meta-analysis of the metrics (C-statistics and accuracy) was performed using R4.2.0 (R development Core Team, Vienna, http://www.R-project.org). If C-statistics lacked

95% confidence interval (CI) and standard error (SE), the SE was estimated according to the method by Debray TP et al. [26]. For inaccurate original data, we conducted calculations based on sensitivity and specificity as well as the sample size of each molecular subtype and the sample size of modeling. Given the difference in variables and parameters among ML models, the random-effects model was utilized for meta-analysis.

Results

Literature screening

A comprehensive and systematic search was first conducted in Pubmed, Cochrane, Embase, and Web of Science until May 25, 2023. A total of 2, 820 articles were retrieved. After removing 491 duplicates, we reviewed the titles and abstracts of the remaining 2, 329 papers. The full texts of 16 potentially eligible studies were downloaded. After a full-text review, 1 unpublished conference paper was excluded, 1 paper about prediabetes, not GDM was excluded; and 2 papers were excluded for lacking outcome indicators to measure model accuracy. Therefore, 12 original studies were included in the first stage.

In addition, we conducted a supplemental search on July 2, 2024, with 443 publications retrieved from various databases, and 1 newly published study was added after the same screening process. Eventually, a total of 13 studies [27–39] were included in the present research (Fig. 1).

Characteristics of included studies

A total of 13 eligible studies were included in the present study, involving 11,320 patients with a history of GDM. During the follow-up period, 1,246 cases progressed to T2DM (Tables 1 and 2). The specific characteristics of the 13 eligible studies are shown in Fig. 1. These studies were published from 2011 to 2024 in eight countries, including Canada, China, Germany, the United States, the United Kingdom, Australia, Italy, and Israel. There were 6 prospective studies [27, 28, 30, 32, 35, 39], 1 randomized controlled trial [34], and 6 retrospective studies [29, 31, 33, 36–38]. Among these studies, two studies [28, 35] conducted external validation, and five studies performed cross-validation [27, 31, 33, 34, 39]. The mean follow-up duration was 5 years.

Quality assessment

The ROB of the studies included was assessed using PROBAST. The model types included Logistic Regression (LR), Decision Trees (DT), Naive Bayes (NB), Cox, and Random Forest (RF), among which Cox and DT were the predominant types. Of the studies we included, 7 models came from case-control studies, and 2 models came from studies that did not specify the type, so the risk of bias was unclear. 7 models from case-control studies had a high risk of bias in predictor evaluation, and 1 model contained a large number of missing values, and therefore it had a high risk of bias in predictors. The number

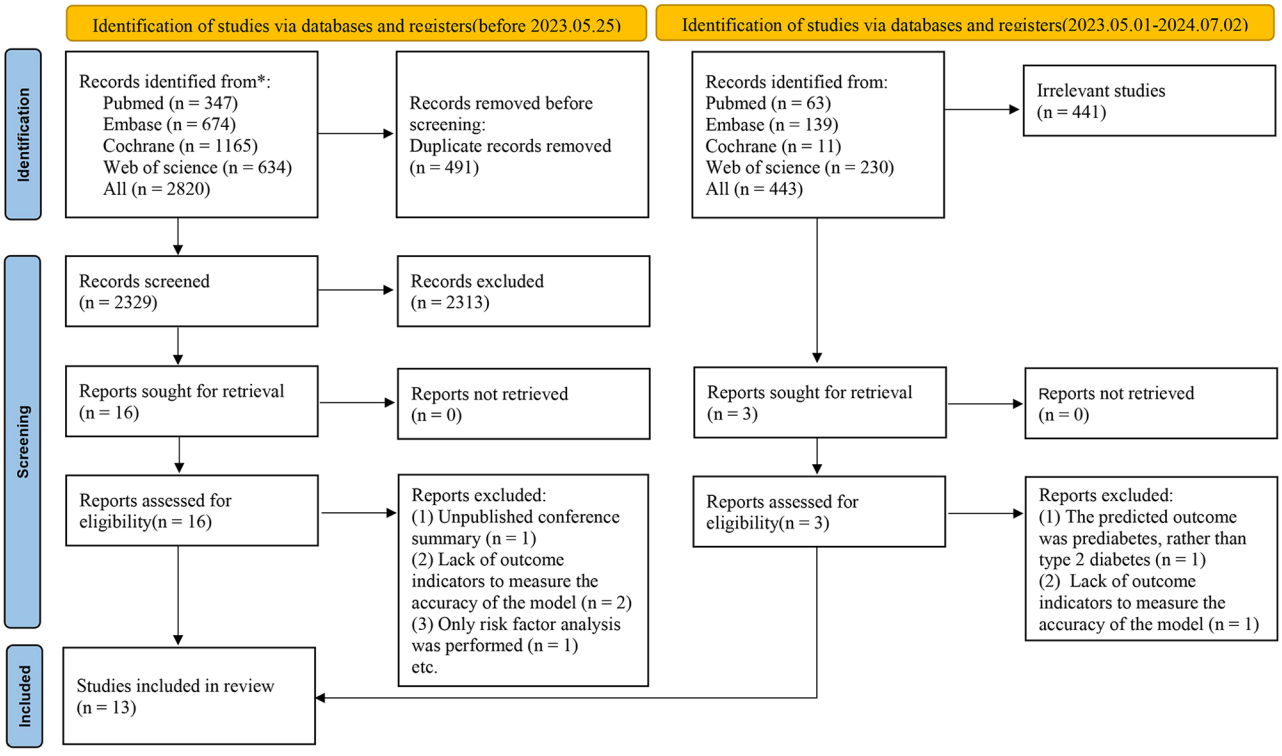

Fig. 1 Flow diagram of study selection

Table 1 Basic characteristics of the included studies

No.	Author	Year	Country of author	Study type	Number of type 2 diabetes samples in training set	Sample size of training set	Number of type 2 diabetes samples in verification set	Sample size of verification set	Follow-up duration
1	Saifur R Khan [27]	2019	Canada	Prospective cohort	55	140			2 years
2	Amina Allalou [28]	2016	Canada	Prospective cohort	80	160	42	84	2 years
3	Weiqin Li [29]	2020	China	Retrospective cohort	83	1263			2.3 years
4	Mi Lai [30]	2020	Canada	Prospective cohort	122	556			2 years
5	Hung-Chun Lin [31]	2011	China taiwan	Retrospective cohort	152	558			4 years
6	Köhler M [32]	2016	Germany	Prospective cohort	110	257			20 years
7	Nishanthi Periyathambi [33]	2022	UK	Retrospective cohort	11	394			2 years
8	Bernice Man [34]	2021	USA	RCT	82	317			3.2 years
9	Mugdha V Joglekar [35]	2021	Australia	Prospective cohort	21	103	11	82	10 years
10	*Ludovica Ilari* [36]	2022	Italy	Retrospective cohort	19	78			7 years
11	Ohad Houri [37]	2022	Israel	Retrospective cohort	366	6091			5.3 years
12	Martha Lappas [38]	2015	Australia	Retrospective cohort	21	104			8.5 years
13	Yitayeh Belsti [39]	2024	Australia	Prospective cohort	124	1299			3 years

of missing values could not be determined for 1 model; therefore, the validity of its predictors was unclear. 6 models had an unclear risk of bias in outcomes since it was unable to determine whether the information about the predictors was clear when assessing the outcome. In addition, 17 models were not satisfied with an EPV > 20 or did not have an independent validation set, and therefore they had a high risk of bias in statistical analysis (Fig. 2). Model variables encompassed biomarkers of metabolic disease, lipolytic metabolites, and independent predictors of T2DM.

Results of meta-analysis

C-statistic

23 ML models reported the C-statistic and a random-effects model was used for meta-analysis. The pooled C-statistic was 0.82 (95% CI: 0.79 ~ 0.86) (Fig. 3). A subgroup analysis by model type was performed, which yielded a pooled C-statistic of 0.81 (95% CI: 0.75 ~ 0.88) for LR, 0.85 (95% CI: 0.80 ~ 0.90) for DT, 0.79 (95% CI: 0.71 ~ 0.86) for NB, 0.80 (95% CI: 0.74 ~ 0.85) for Cox, and 0.88 (95% CI: 0.81 ~ 0.94) for RF.

Sensitivity and specificity

Due to a significant imbalance in the number of patients with and without T2DM in original studies, the C-statistics of ML models could not adequately explain the accuracy in predicting T2DM. Therefore, meta-analyses of sensitivity and specificity were conducted for the 23 ML models (Fig. 4), which yielded a pooled sensitivity of 0.76 (0.72–0.80) and a pooled specificity of 0.57 (0.50–0.65).

Discussion

This paper summarized relevant studies published in recent years. The overall C-statistic was 0.82 (95%CI: 0.79 ~ 0.86), indicating the favorable performance of ML models in the prediction of the risk of progressing to T2DM in women with a history of GDM, because a C-statistic > 0.7 generally demonstrated an ideal result [40]. High heterogeneity was observed across eligible studies due to the diversity of ML models, mainly in ML algorithms, modeling indicators, risk indicators, and modeling parameters.

The 75-g oral glucose tolerance test (OGTT) and fasting plasma glucose (FPG) tests were recommended by the American Diabetes Association (ADA) for screening T2DM at the 6 to 12 weeks postpartum and every 1–3 years thereafter [41]. Another study also recommended that women at high risk of T2DM should be evaluated for A1C levels [42]. Despite these recommendations, screening of GDM after pregnancy remains notably inadequate, with very low compliance rates [28, 43], though the integrated healthcare system reported a screening rate of 60% [44]. The low screening rate is attributed to

Table 2 Modeling information of the included studies

No	Author	Year	Variable selection method	Model type	Modeling variables
1	Saifur R Khan [27]	2019	Stepwise (bidirectional) multiple logistic regression	Logistic regression, decision tree	FPG,2-hour glucose and family history of diabetes, CE17:0,NEFA22:5,LPE20:4,TAG47:0-FA16:0,TAG48:1-FA14:0,TAG49:0-FA16:0,TAG52:1-FA18:0,TAG52:5-FA20:5,TAG52:6-FA22:6,TAG54:1-FA16:0,TAG54:3-FA20:3,TAG54:5-FA22:5 (lipid metabolites) -------LR CE, NEFA, TAG, FA, PE (lipid metabolites)------decision tree
2	Amina Allalou [28]	2016	Model itself	Decision tree, Naive Bayes	PC ae C40:5/BCAAs/Hexos/SM(OH) (metabolite) PC ae C40:5/BCAAs/Hexos/SM(OH) (metabolite) + metabolin
3	Weiqin Li [29]	2020	Multivariate Cox proportional hazards regression	Cox	Family history of diabetes, pre-pregnant body mass index [BMI], gestational age, gestational hypertension, 2-hour 75 g OGTT at 26–30 gestational weeks, postpartum T2D ------ Multivariate Cox proportional hazards regression model
4	Mi Lai [30]	2020	Random forest	Random forest	Hexose, His, Spermidine, AC10, Kynurenine, lysoPC a C26:0, total DMA, PC ae C40:4, Ser, Glu, PC aa C32:2, Tyr, PC aa C30:0, Ile, AC3, lysoPC a C26:1, SM(OH)C22:2, Gln, PC aa C32:1, SM C20:1 (metabolite)
5	Hung-Chun Lin [31]	2011	LASSO	Artificial immune system	Motherly age,50-g oral glucose challenge test value, Fasting glucose value,1-h 100-g OGTT value, 2-h 100-g OGTT value,3-h 100-g OGTT value, BMI before gestation, Weight before gestation, Increased weight during pregnancy, Weight of newborn, Family history of diabetes mellitus, The number of pregnancy times
6	Köhler M [32]	2016	LASSO	LASSO	the mother's body mass index (BMI) in early, pregnancy, GDM treatment (insulin or diet), family history of diabetes (yes vs. no), lactation duration (never, ≤ 3 months, or > 3 months, coded as two dummy variables), and maternal age at delivery
7	Nishanthi Periyathambi [33]	2022	Univariate + multivariate logistic	Logistic	Intercept, Maternal age, Antenatal fasting glucose, Antenatal 2-hrs Glucose, Antenatal HbA1c, Gestational age at GDM diagnosis, Booking BMI, Continuing to smoke at booking, Unmarried at booking, Diastolic BP at booking, Other Ethnicity (Black African/ Caribbean or mixed ethnicity), Gestational age at birth, Instrument assisted delivery, Women delivered SGA infants, Women delivered male babies, Breastfeeding initiation before discharge
8	Bernice Man [34]	2021	Univariate + multivariate Cox	Cox	age group, ethnicity, parental (either mother or father), history of diabetes (type not specified), BMI group, waist circumference, waist-to-hip ratio, fasting glucose and triglycerides, HbA1c, self-reported physical activity, and treatment arm(placebo, ILI, or metformin)
9	Mugdha V Joglekar [35]	2021	LASSO	Cox	Age, BMI, Pregnancy fasting glucose, Postnatal fasting glucose, Cholesterol, Triacylglycerol, miR-491-5p, miR-543, miR-410-3p, miR-410-3p, miR-491-5p + miR-543 + miR-410-3p + miR-369-3p, Base model (six traditional risk factors combined), Base model + miR-369-3p (biomarkers)
10	*Ludovica Ilari* [36]	2022	Decision tree, Naive Bayes, and logistic regression	Decision tree, Naive Bayes, and logistic regression	Age, BW(Body weight), h(Height), BMI(Body mass index), gb(Basal glucose), GMEAN(Mean area under the glucose curve), IMEAN(Mean area under the insulin curve), CpMEAN(Mean area under the C-peptide curve), AUCINS–1P (Area under the insulin curve during the 1st phase of test), AUCINS–2P(Area under the insulin curve during the 2nd phase of test), KG1(Disappearance rate of glucose before insulin injection), KG2(Disappearance rate of glucose after insulin injection), SI(Insulin sensitivity), SG(Glucose effectiveness), V(Distribution volume of glucose), BIE(Basal insulin effect of glucose effectiveness), GEZI(Glucose effectiveness at zero insulin), AIR(Mean of suprabasal insulin in the time interval 3–8 min), ACPR(Mean of suprabasal C-peptide in the time interval 3–8 min), DI(Disposition index), BSR(Basal secretion rate), Φ1cz(β-cell responsivity to glucose), AUCSECR(Area under the secretion curve during the entire test), AUCSECR–1P(Area under the secretion curve during the 1st phase of test), AUCSECR–2P(Area under the secretion curve during the 2nd phase of test), CLMEAN(Mean insulin clearance during the entire test), CLMEAN–1P(Mean insulin clearance during the 1st phase of test), CLMEAN–2P(Mean insulin clearance during the 2nd phase of test), CLp(Extra-hepatic insulin clearance), FEL(Hepatic insulin clearance), IPEAK–FIRST(Peak insulin after glucose injection), IPEAK–INJECT(Peak insulin after insulin injection), CPEAK(Peak C-peptide), DOSE(Glucose dose injected)
11	Ohad Houri [37]	2022	LASSO	Decision tree, Lasso Cox	age, parity, gravidity, oral glucose tolerance test results—both the 1-h, 50-g, glucose challenge test (GCT) and the 3-h, 100-g, oral glucose tolerance test (OGTT), gestational age at delivery and birthweight

Table 2 (continued)

No	Author	Year	Variable selection method	Model type	Modeling variables
12	Martha Lappas [38]	2015	Univariate + multivariate logistic	Logistic	Age, BMI, Pregnancy fasting OGTT, Postnatal fasting OGTT, three lipid species: PE(P-36:2)(alkenylphosphatidylethanolamine species), PS 38:4(phosphatidylserine species), CE 20:4(cholesteryl ester species)
13	Yitayeh Belsti [39]	2024	Logistic	Logistic	antenatal FPG, antenatal 2 h-OGTT, history of recurrent GDM, insulin treatment during pregnancy, parity, history of irregular menstrual cycle, and family history of diabetes mellitus, antenatal 2 h-OGTT, postnatal 2 h-OGTT, postnatal FPG, and BMI

Note: cox-proportional hazards model

LASSO-Least absolute shrinkage and selection operator

Logistic-logistic regression

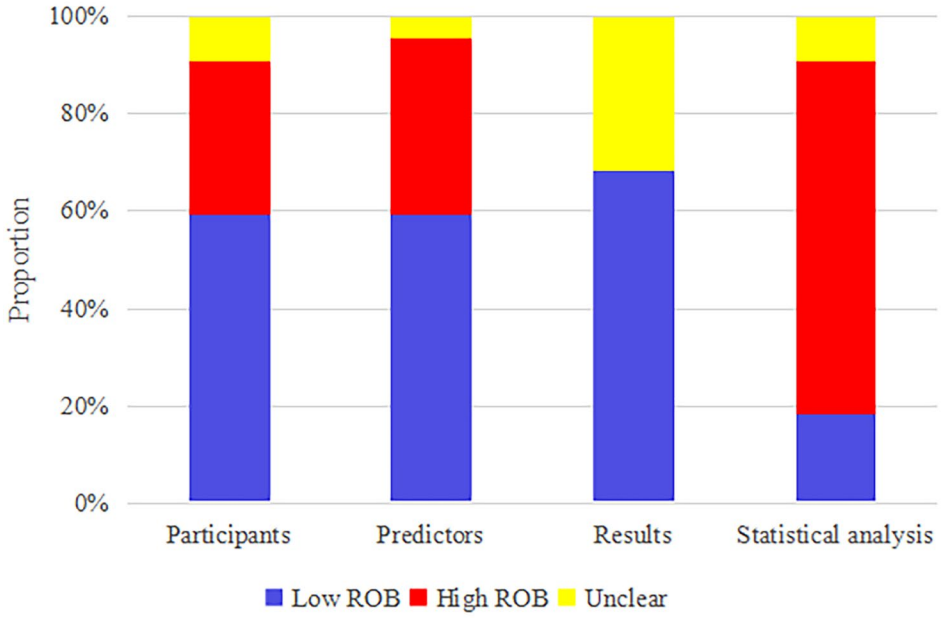

Fig. 2 Quality assessment results of ML models in included studies

the difficulties in performing OGTT, patients' fear of being diagnosed with diabetes, and lack of screening at postpartum visits [45]. Existing diagnostic tests, such as FPG, 2-hour plasma glucose, and HbA1c, have performance limitations in identifying high-risk populations [46]. Women with a history of GDM are seven times more likely to develop T2DM in later life [29]. OGTT is currently used to assess the risk of T2DM in women with a history of GDM, but it has many limitations. There is a lack of accurate risk stratification methods for postpartum T2DM. Risk prediction models may aid in clinical decision-making for patient management [47]. In recent years, ML has garnered much attention [48]. ML, a subset of artificial intelligence, focuses on developing computer systems that can identify potential patterns in training data to perform classification and prediction tasks on new data. ML integrates tools from statistics, data mining, and optimization to create models [49]. It has potentially ideal value and exhibits superior performance to OGTT.

Among ML algorithms (LR, DT, NB, Cox, and RF) in the included studies, Cox and DT were the most utilized. No significant difference was observed in the C-statistics of these ML algorithms. LR exhibited inferior predictive accuracy to other algorithms. Nevertheless, the lack of significant differences in their predictive performance suggests that further investigation is warranted to figure out more effective predictors in future analyses.

Variables

Variables included lipolytic metabolites, independent predictors of T2DM, and biomarkers of metabolic diseases. LR and DT were predominantly used among ML models. Two eligible studies had the highest C-statistics. ML optimization in a DT format revealed a T2DM predictive signature [27] with a C-statistic of 0.92 and an

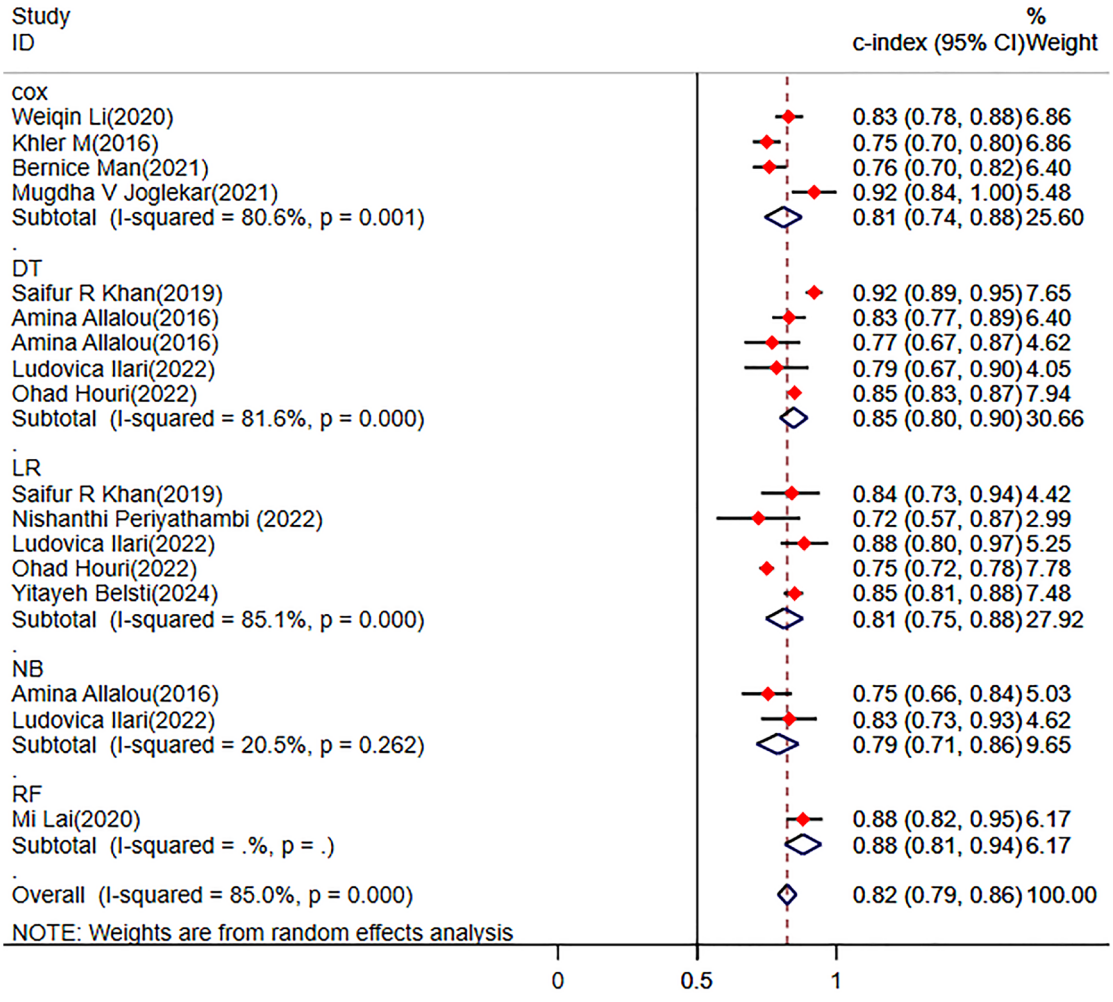

Fig. 3 Forest plot of C-statistic

accuracy of 91%. The accuracy of this predictive signature was not affected by sensitivity (87%) or specificity (93%) and was more accurate than clinical diagnoses. Circulating microRNAs [50] are well-established biomarkers for metabolic diseases. The inclusion of circulating miR-369-3p measured at 12 weeks postpartum significantly increased the C-statistic from 0.83 (95% CI 0.68, 0.97) to 0.92 (95% CI: 0.84, 1.00) for predicting future T2DM in women with previous GDM [35]. Incorporating lipid metabolites and biomarkers of metabolic diseases into the prediction model for the progression of GDM to T2DM provides greater prediction accuracy for future risk [51].

Advantages and limitations

This paper first summarized the predictive accuracy of ML for the risk of T2DM in women with a history of GDM and provided evidence for the development and updating of subsequent clinical prediction tools. Nevertheless, there are some limitations. First, a few studies and models were included in the present paper after a comprehensive and systematic review. Second, many included studies were from single centers and lacked independent external validation. Only a few of the studies included had external validation, which may limit the interpretation of our results. Third, some studies included had fewer cases. Fourth, the follow-up duration of the studies included varied significantly. Due to the small number of models included, we did not discuss the differences in follow-up durations in depth under the same model. Heterogeneity may be introduced due to varied models, outcome measures, and follow-up durations, which is a limitation of our study.

Conclusion

ML has favorable predictive accuracy for the progression to T2DM in women with a history of GDM and can be used as a potential predictive tool. However, this conclusion is based on limited evidence. Therefore, real-time prediction models with broader applicability should be

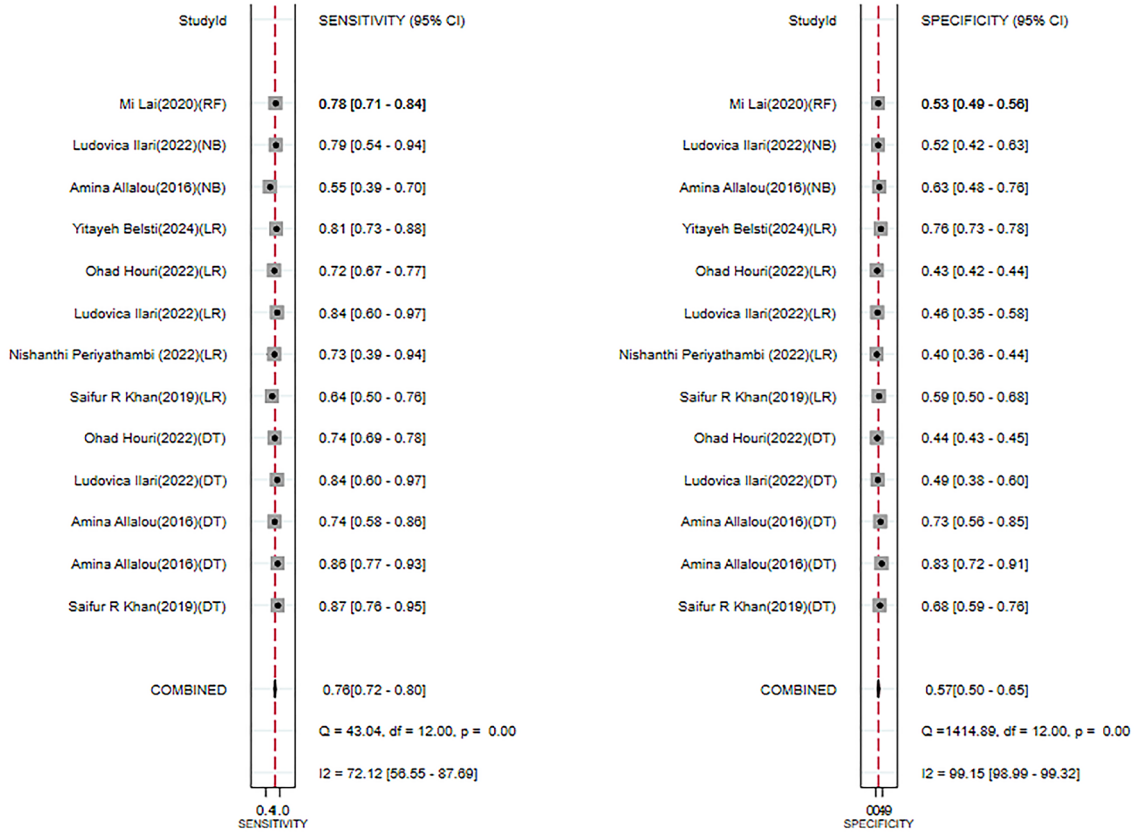

Fig. 4 Forest plot of sensitivity and specificity

developed for GDM patients of different ethnic backgrounds. Individualized health interventions and lifestyle recommendations for high-risk populations will be beneficial to reduce the risk of T2DM.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-024-02848-x.

Supplementary Material 1: Table S1 Search strategy

Supplementary Material 2

Acknowledgements
Not applicable.

Author contributions
Yijun Xu, Lili Wei, and Meng Zhao conceived the idea of this study; Meng Zhao and Lili Wei screened the literature and extracted the data; Meng Zhao conducted statistical analysis; Lidan Ma supervised the analysis. Meng Zhao, Zhixin Yao, Yan Zhang, Wenquan Pang, and Shuyin Ma explained the survey results; Meng Zhao drafted the manuscript. Yijun Xu, Lili Wei, Meng Zhao, and Lidan Ma critically reviewed the manuscript, and Meng Zhao revised the manuscript for final submission. All authors have approved the final draft of the manuscript. Yijun Xu is the guarantor. Meng Zhao takes full responsibility for the study and has the right to access data and decide the publication. The corresponding author certified that all listed authors met the author criteria, and no other authors who met the criteria were omitted.

Funding
Quality of life care research fund of Chinese Association For Life Care (HL20210145).

Data availability
The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 30 October 2023 / Accepted: 31 December 2024
Published online: 13 January 2025

References

1. Wang J, Zheng J, Shi W, Du N, Xu X, Zhang Y, et al. Dysbiosis of maternal and neonatal microbiota associated with gestational diabetes mellitus. Gut. 2018;67(9):1614–25. https://doi.org/10.1136/gutjnl-2018-315988.
2. Saravanan P. Gestational diabetes: opportunities for improving maternal and child health. Lancet Diabetes Endocrinol. 2020;8(9):793–800. https://doi.org/10.1016/s2213-8587(20)30161-3.

3. Ravaut M, Harish V, Sadeghi H, Leung KK, Volkovs M, Kornas K, et al. Development and validation of a Machine Learning Model Using Administrative Health Data to predict onset of type 2 diabetes. JAMA Netw Open. 2021;4(5):e2111315. https://doi.org/10.1001/jamanetworkopen.2021.11315.
4. 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:m1361. https://doi.org/10.1136/bmj.m1361.
5. Ratner RE, Christophi CA, Metzger BE, Dabelea D, Bennett PH, Pi-Sunyer X, et al. Prevention of diabetes in women with a history of gestational diabetes: effects of metformin and lifestyle interventions. J Clin Endocrinol Metab. 2008;93(12):4774–9. https://doi.org/10.1210/jc.2008-0772.
6. Green JB. Cardiovascular consequences of Gestational Diabetes. Circulation. 2021;143(10):988–90. https://doi.org/10.1161/circulationaha.120.052995.
7. 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. https://doi.org/10.1007/s11892-015-0699-x.
8. Kim C, Berger DK, Chamany S. Recurrence of gestational diabetes mellitus: a systematic review. Diabetes Care. 2007;30(5):1314–9. https://doi.org/10.2337/dc06-2517.
9. Khosla S, Samakkarnthai P, Monroe DG, Farr JN. Update on the pathogenesis and treatment of skeletal fragility in type 2 diabetes mellitus. Nat Rev Endocrinol. 2021;17(11):685–97. https://doi.org/10.1038/s41574-021-00555-5.
10. Segar MW, Vaduganathan M, Patel KV, McGuire DK, Butler J, Fonarow GC, et al. Machine learning to predict the risk of Incident Heart failure hospitalization among patients with diabetes: the WATCH-DM risk score. Diabetes Care. 2019;42(12):2298–306. https://doi.org/10.2337/dc19-0587.
11. Carson MP, Ananth CV, Gyamfi-Bannerman C, Smulian J, Wapner RJ. Postpartum Testing to Detect Persistent Dysglycemia in Women with Gestational Diabetes Mellitus. Obstet Gynecol. 2018;132(1):193–8. https://doi.org/10.1097/aog.0000000000002687.
12. Hod M, Kapur A, McIntyre HD. Evidence in support of the International Association of Diabetes in pregnancy study groups' criteria for diagnosing gestational diabetes mellitus worldwide in 2019. Am J Obstet Gynecol. 2019;221(2):109–16. https://doi.org/10.1016/j.ajog.2019.01.206.
13. Chen X, Zhang Y, Chen H, Jiang Y, Wang Y, Wang D, et al. Association of Maternal Folate and Vitamin B(12) in early pregnancy with gestational diabetes Mellitus: a prospective cohort study. Diabetes Care. 2021;44(1):217–23. https://doi.org/10.2337/dc20-1607.
14. Salvia MG, Quatromoni PA. Behavioral approaches to nutrition and eating patterns for managing type 2 diabetes: a review. Am J Med Open. 2023;9:100034. https://doi.org/10.1016/j.ajmo.2023.100034.
15. Kirwan JP, Sacks J, Nieuwoudt S. The essential role of exercise in the management of type 2 diabetes. Cleve Clin J Med. 2017;84(7 Suppl 1):S15–21. https://doi.org/10.3949/ccjm.84.s1.03.
16. AlKhudidi FH, Alsulaimani AI, Alharthi AH, Alrumaym AH, Alharthi EK, Altalhi WA, et al. Awareness of type 2 Diabetic patients about the importance of Exercise and Diet on Diabetes Type 2 in the Western Region of Saudi Arabia. Mater Sociomed. 2021;33(4):276–81. https://doi.org/10.5455/msm.2021.33.276-281.
17. Zhang Z, Yang L, Han W, Wu Y, Zhang L, Gao C, et al. Machine learning prediction models for gestational diabetes Mellitus: Meta-analysis. J Med Internet Res. 2022;24(3):e26634. https://doi.org/10.2196/26634.
18. Andaur Navarro CL, Damen JAA, Takada T, Nijman SWJ, Dhiman P, Ma J, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ. 2021;375:n2281. https://doi.org/10.1136/bmj.n2281.
19. Bedrikovetski S, Dudi-Venkata NN, Kroon HM, Seow W, Vather R, Carneiro G, et al. Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. BMC Cancer. 2021;21(1):1058. https://doi.org/10.1186/s12885-021-08773-w.
20. Bedrikovetski S, Dudi-Venkata NN, Maicas G, Kroon HM, Seow W, Carneiro G, et al. Artificial intelligence for the diagnosis of lymph node metastases in patients with abdominopelvic malignancy: a systematic review and meta-analysis. Artif Intell Med. 2021;113:102022. https://doi.org/10.1016/j.artmed.2021.102022.
21. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Med. 2021;13(1):152. https://doi.org/10.1186/s13073-021-00968-x.
22. Yang Y, Xu L, Sun L, Zhang P, Farid SS. Machine learning application in personalised lung cancer recurrence and survivability prediction. Computational and Structural Biotechnology Journal. 2022; 20:1811–20. https://doi.org/10.1016/j.csbj.2022.03.035
23. Liang X, Yu X, Gao T. Machine learning with magnetic resonance imaging for prediction of response to neoadjuvant chemotherapy in breast cancer: a systematic review and meta-analysis. Eur J Radiol. 2022;150:110247. https://doi.org/10.1016/j.ejrad.2022.110247.
24. Vieira S, Liang X, Guiomar R, Mechelli A. Can we predict who will benefit from cognitive-behavioural therapy? A systematic review and meta-analysis of machine learning studies. Clin Psychol Rev. 2022;97:102193. https://doi.org/10.1016/j.cpr.2022.102193.
25. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100.
26. Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. Stat Methods Med Res. 2019;28(9):2768–86. https://doi.org/10.1177/0962280218785504.
27. Khan SR, Mohan H, Liu Y, Batchuluun B, Gohil H, Al Rijjal D, et al. The discovery of novel predictive biomarkers and early-stage pathophysiology for the transition from gestational diabetes to type 2 diabetes. Diabetologia. 2019;62(4):687–703. https://doi.org/10.1007/s00125-018-4800-2.
28. Allalou A, Nalla A, Prentice KJ, Liu Y, Zhang M, Dai FF, et al. A predictive metabolic signature for the Transition from Gestational Diabetes Mellitus to type 2 diabetes. Diabetes. 2016;65(9):2529–39. https://doi.org/10.2337/db15-1720.
29. Li W, Leng J, Liu H, Zhang S, Wang L, Hu G, et al. Nomograms for incident risk of post-partum type 2 diabetes in Chinese women with prior gestational diabetes mellitus. Clin Endocrinol. 2019;90(3):417–24. https://doi.org/10.1111/cen.13863.
30. Lai M, Liu Y, Ronnett GV, Wu A, Cox BJ, Dai FF, et al. Amino acid and lipid metabolism in post-gestational diabetes and progression to type 2 diabetes: a metabolic profiling study. PLoS Med. 2020;17(5):e1003112. https://doi.org/10.1371/journal.pmed.1003112.
31. Lin HC, Su CT, Wang PC. An application of artificial immune recognition system for prediction of diabetes following gestational diabetes. J Med Syst. 2011;35(3):283–9. https://doi.org/10.1007/s10916-009-9364-8.
32. Köhler M, Ziegler AG, Beyerlein A. Development of a simple tool to predict the risk of postpartum diabetes in women with gestational diabetes mellitus. Acta Diabetol. 2016;53(3):433–7. https://doi.org/10.1007/s00592-015-0814-0.
33. Periyathambi N, Parkhi D, Ghebremichael-Weldeselassie Y, Patel V, Sukumar N, Siddharthan R, et al. Machine learning prediction of non-attendance to postpartum glucose screening and subsequent risk of type 2 diabetes following gestational diabetes. PLoS ONE. 2022;17(3):e0264648. https://doi.org/10.1371/journal.pone.0264648.
34. Man B, Schwartz A, Pugach O, Xia Y, Gerber B. A clinical diabetes risk prediction model for prediabetic women with prior gestational diabetes. PLoS ONE. 2021;16(6):e0252501. https://doi.org/10.1371/journal.pone.0252501.
35. Joglekar MV, Wong WKM, Ema FK, Georgiou HM, Shub A, Hardikar AA, et al. Postpartum circulating microRNA enhances prediction of future type 2 diabetes in women with previous gestational diabetes. Diabetologia. 2021;64(7):1516–26. https://doi.org/10.1007/s00125-021-05429-z.
36. Ilari L, Piersanti A, Göbl C, Burattini L, Kautzky-Willer A, Tura A, et al. Unraveling the factors determining development of type 2 diabetes in Women with a history of gestational diabetes Mellitus through Machine-Learning techniques. Front Physiol. 2022;13:789219. https://doi.org/10.3389/fphys.2022.789219.
37. Houri O, Gil Y, Chen R, Wiznitzer A, Hochberg A, Hadar E et al. Prediction of type 2 diabetes Mellitus according to glucose metabolism patterns in pregnancy using a Novel Machine Learning Algorithm. J J Med Biol Eng. 2022.
38. Lappas M, Mundra PA, Wong G, Huynh K, Jinks D, Georgiou HM, et al. The prediction of type 2 diabetes in women with previous gestational diabetes mellitus using lipidomics. Diabetologia. 2015;58(7):1436–42. https://doi.org/10.1007/s00125-015-3587-7.
39. Belsti Y, Moran LJ, Goldstein R, Mousa A, Cooray SD, Baker S, et al. Development of a risk prediction model for postpartum onset of type 2 diabetes mellitus, following gestational diabetes; the lifestyle InterVention in gestational diabetes (LIVING) study. Clin Nutr. 2024;43(8):1728–35. https://doi.org/10.1016/j.clnu.2024.06.006.
40. Jiang W, Wang J, Shen X, Lu W, Wang Y, Li W, et al. Establishment and validation of a risk Prediction Model for Early Diabetic kidney Disease based

on a Systematic Review and Meta-analysis of 20 cohorts. Diabetes Care. 2020;43(4):925–33. https://doi.org/10.2337/dc19-1897.
41. 12. Management of Diabetes in Pregnancy. Diabetes Care. 2016;39(Suppl 1):S94–8. https://doi.org/10.2337/dc16-S015.
42. Standards of medical care in diabetes–2014. Diabetes Care. 2014;37(Suppl 1):S14–80. https://doi.org/10.2337/dc14-S014.
43. Blatt AJ, Nakamoto JM, Kaufman HW. Gaps in diabetes screening during pregnancy and postpartum. Obstet Gynecol. 2011;117(1):61–8. https://doi.org/10.1097/AOG.0b013e3181fe424b.
44. Ferrara A. Increasing prevalence of gestational diabetes mellitus: a public health perspective. Diabetes Care. 2007;30(Suppl 2):S141–6. https://doi.org/10.2337/dc07-s206.
45. Russell MA, Phipps MG, Olson CL, Welch HG, Carpenter MW. Rates of postpartum glucose testing after gestational diabetes mellitus. Obstet Gynecol. 2006;108(6):1456–62. https://doi.org/10.1097/01.Aog.0000245446.85868.73.
46. Jagannathan R, Neves JS, Dorcely B, Chung ST, Tamura K, Rhee M, et al. The oral glucose tolerance test: 100 years later. Diabetes Metab Syndr Obes. 2020;13:3787–805. https://doi.org/10.2147/dmso.S246062.
47. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC Med. 2011;9:103. https://doi.org/10.1186/1741-7015-9-103.
48. Kodama S, Fujihara K, Horikawa C, Kitazawa M, Iwanaga M, Kato K, et al. Predictive ability of current machine learning algorithms for type 2 diabetes mellitus: a meta-analysis. J Diabetes Investig. 2022;13(5):900–8. https://doi.org/10.1111/jdi.13736.
49. Fregoso-Aparicio L, Noguez J, Montesinos L, García-García JA. Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. Diabetol Metab Syndr. 2021;13(1):148. https://doi.org/10.1186/s13098-021-00767-9.
50. Capobianco E. Systems and precision medicine approaches to diabetes heterogeneity: a Big Data perspective. Clin Transl Med. 2017;6(1):23. https://doi.org/10.1186/s40169-017-0155-4.
51. Scirica BM. Use of biomarkers in Predicting the Onset, Monitoring the Progression, and risk stratification for patients with type 2 diabetes Mellitus. Clin Chem. 2017;63(1):186–95. https://doi.org/10.1373/clinchem.2016.255539.

Publisher's note