111309_Revision_Auto_Edited.docx

WORD COUNT	6067	TIME SUBMITTED	03-AUG-2025 08:42AM
		PAPER ID	117544536

Name of Journal: *World Journal of Diabetes*
Manuscript NO: 111309
Manuscript Type: REVIEW

Advances in gestational diabetes mellitus screening: Emerging trends and future directions

Didem Kaymak, A. Seval Ozgu-Erdinc

Abstract

Gestational diabetes mellitus (GDM) is a multifactorial metabolic disorder first recognized during pregnancy, with rising global prevalence and significant implications for both maternal and neonatal outcomes. This review provides a comprehensive synthesis of current diagnostic strategies, including standard screening protocols such as the one-step and two-step oral glucose tolerance tests, and evaluates their limitations in terms of sensitivity, timing, and practicality. The complex pathogenesis of GDM—centered on β-cell dysfunction, insulin resistance, adipose tissue dysregulation, placental transport abnormalities, and neurohormonal imbalance—is explored in detail, highlighting the interplay of metabolic, inflammatory, and epigenetic mechanisms. Particular emphasis is placed on the emerging role of predictive biomarkers, encompassing metabolic, inflammatory, placental, urinary, and genetic indicators. These biomarkers, including adipokines, angiogenic factors, and microRNAs, offer promising avenues for early identification of at-risk individuals prior to the onset of hyperglycemia. The review also assesses recent advances in machine learning–based risk prediction models, which have demonstrated superior accuracy over traditional algorithms and may facilitate personalized screening and management strategies. Despite encouraging findings, challenges such as biomarker standardization,

ethnic variability, and model validation persist. This review underscores the necessity for integrated, multi-omic, and patient-centered approaches to optimize GDM prediction, early diagnosis, and long-term risk reduction for both mother and child.

Key Words: Gestational diabetes mellitus; predictive biomarkers; insulin resistance; β-cell dysfunction; machine learning; early screening

Core Tip: This review offers an integrative and up-to-date overview of GDM, emphasizing the evolving understanding of its pathophysiology and the emerging role of multi-dimensional biomarkers in early prediction. By combining evidence on metabolic, inflammatory, placental, genetic, and urinary biomarkers, alongside advanced machine learning–based models, this work underscores the shift toward precision diagnostics. It critically evaluates conventional screening strategies and highlights avenues for improving early detection and individualized care. The synthesis aims to support clinicians and researchers in refining GDM risk stratification and mitigating long-term maternal-fetal consequences.

INTRODUCTION

Gestational diabetes mellitus (GDM) is a transient yet clinically significant metabolic disorder first recognized during pregnancy, characterized by glucose intolerance leading to hyperglycemia (1). The condition arises due to pancreatic β-cell dysfunction combined with increased insulin resistance, driven by hormonal changes during pregnancy (2). Insulin resistance progressively worsens as gestation advances, often peaking in the third trimester, making early identification and intervention crucial (1). DM prevalence varies globally, ranging from 6% to 20%, influenced by ethnic, genetic, and environmental factors, with an increasing incidence due to the obesity epidemic and changing diagnostic criteria (3, 4). The incidence of GDM increases with the presence of risk factors common to type 2 diabetes, such as obesity and advanced maternal age. The risk factors are shown in Box 1 (5). Women diagnosed

with GDM face significant maternal and neonatal risks, including preeclampsia, fetal macrosomia, neonatal hypoglycemia, and a heightened lifetime risk of type 2 diabetes mellitus (T2DM). Additionally, long-term consequences for offspring include childhood obesity and an increased risk of metabolic syndrome (Box 2) (6, 7).

Standard screening for GDM typically involves an oral glucose tolerance test (OGTT) between 24 and 28 weeks of gestation, following a 50-g glucose challenge test (GCT) in some protocols. While both methods are widely accepted, regional and institutional preferences vary based on population risk profiles, accessibility, and healthcare system resources. However, limitations exist in both strategies, including suboptimal predictive capacity for early intervention, variability in patient adherence, and the inconvenience of prolonged fasting and glucose load administration (5, 8). Consequently, alternative biomarkers and technological innovations, such as machine learning (ML)-based risk assessment models, have been explored to refine early screening and diagnostic accuracy (9, 10).

STANDART SCREENING METHODS FOR GDM

According to the American Diabetes Association (ADA) 2025 guidelines, the one-step 75-g OGTT is the preferred method, requiring a fasting glucose measurement, followed by 1-hour and 2-hour glucose readings (5). The ADA defines GDM as present if any of the following thresholds are met: Fasting glucose: ≥ 92 mg/dL (5.1 mmol/L), 1-hour glucose: ≥ 180 mg/dL (10.0 mmol/L), 2-hour glucose: ≥ 153 mg/dL (8.5 mmol/L) (5). Alternatively, the two-step approach, recommended by the American College of Obstetricians and Gynecologists (ACOG) 2018 guidelines, begins with a 50-g GCT (non-fasting). If the 1-hour plasma glucose level is ≥ 140 mg/dL (7.8 mmol/L), a 100-g OGTT is conducted after an overnight fast. GDM is diagnosed if at least two of the following criteria are met: Fasting glucose: ≥ 95 mg/dL (5.3 mmol/L), 1-hour glucose: ≥ 180 mg/dL (10.0 mmol/L), 2-hour glucose: ≥ 155 mg/dL (8.6 mmol/L), 3-hour glucose: ≥ 140 mg/dL (7.8 mmol/L) (Box 3) (8).

Screening during the preconception period is recommended for individuals at risk of diabetes or those belonging to high-risk populations (11-13). Preconceptional screening facilitates the management of pre-existing diabetes, enabling the achievement of lower HbA1c levels prior to pregnancy. This, in turn, has been associated with a reduced incidence of congenital anomalies, preterm birth, perinatal mortality, and neonatal intensive care unit (NICU) admissions (8). Among individuals with risk factors (Box 1) who have not undergone preconceptional screening, early universal screening before the 15th gestational week has been reported to be a more effective approach compared to selective screening (13-15). In individuals diagnosed with GDM in early pregnancy, the risks of preeclampsia, macrosomia, shoulder dystocia, and perinatal mortality are increased, along with a higher likelihood of requiring insulin therapy (16-18). Conversely, no significant association has been reported between adverse perinatal outcomes and HbA1c levels below 5.7% (19, 20). Although the effectiveness of treating abnormal glucose metabolism in early pregnancy remains inconclusive, it is recommended that patients receive nutritional counseling and undergo weekly periodic glucose monitoring. In particular, for individuals with fasting blood glucose levels of 110 mg/dL, intensive treatment and close follow-up should be initiated before the 18th gestational week (5). While HbA1c is an inexpensive and widely accessible test, it has been reported to be insufficient in distinguishing between pregestational diabetes and GDM when screening is conducted after the 15th gestational week.

The large-scale, multinational Hyperglycemia and Adverse Pregnancy Outcomes (HAPO) study, which included more than 23,000 pregnant individuals, demonstrated a linear relationship between increasing maternal glycemic levels at 24–28 weeks of gestation and adverse maternal, fetal, and neonatal outcomes. Moreover, the study identified an increased risk of adverse outcomes even within glycemic ranges previously considered normal for pregnancy (21). The International Association of Diabetes and Pregnancy Study Groups (IADPSG), based on the findings of the HAPO study, established diagnostic threshold values for GDM using the one-step 75 g OGTT conducted at 24–28 weeks of gestation. Analysis of adverse pregnancy outcomes

demonstrated a 1.75-fold increased risk in individuals diagnosed with GDM. The adoption of these criteria is projected to increase the prevalence of GDM from 5–6% to 15–20%, raising concerns regarding additional healthcare costs and the potential medicalization of pregnancies previously considered normal (22). Long-term follow-up studies have shown that individuals diagnosed with GDM have a 3.4-fold increased risk of developing prediabetes and type 2 diabetes mellitus postpartum. Furthermore, their offspring exhibit an increased risk of obesity and metabolic disorders.

The gestational diabetes mellitus (GDM) screening strategies recommended by ADA, ACOG, the World Health Organization (WHO), the International Federation of Gynecology and Obstetrics (FIGO), and the National Institute for Health and Care Excellence are presented in Box 4 (5, 8, 23, 24).

PATHOGENESIS OF GDM

The two primary factors contributing to the development of gestational diabetes mellitus (GDM) are β-cell dysfunction and insulin resistance. In most cases, these abnormalities preexist before pregnancy, may progress during gestation, and explain the increased risk of type 2 diabetes mellitus (T2DM) in the postpartum period (Figure 1).

The primary function of β-cells is to synthesize, store, and secrete insulin in response to glucose load. When there is a failure in glucose sensing or insulin secretion, this condition is defined as β-cell dysfunction. It has been proposed that chronic energy excess leads to prolonged and excessive insulin production in β-cells, causing functional impairment (2). However, the pathophysiology of β-cell dysfunction is complex and involves multiple mechanisms (25, 26). Dysfunction may occur at various stages, including proinsulin synthesis, post-translational modifications, insulin granule storage, glucose-sensing mechanisms, or insulin granule exocytosis. The potassium voltage-gated channel KQT-like subfamily member 1 (KCNQ1) and glucokinase (GCK) genes play crucial roles in the regulation of β-cell function. The literature indicates that most genetic variants associated with GDM are linked to β-cell function-related genes.

Additionally, minor defects in β-cell mechanisms may become evident under metabolic stress conditions such as pregnancy, contributing to GDM development (26). β-cell dysfunction is further exacerbated by insulin resistance, leading to reduced glucose uptake and increased β-cell workload, which in turn contributes to hyperglycemia. The direct impact of glucose on β-cell failure is termed glucotoxicity (27). This process establishes a vicious cycle between hyperglycemia, insulin resistance, and β-cell dysfunction. Over time, glucotoxicity triggers β-cell apoptosis, resulting in a decline in pancreatic β-cell mass. In patients with type 2 diabetes mellitus (T2DM), β-cell mass reduction of approximately 40–60% has been reported, and this loss becomes more pronounced as the disease progresses (28, 29). Based on existing animal and human studies, it has been proposed that a combination of β-cell mass reduction, functional impairment, and inadequate insulin secretion contributes to the development of GDM (30).

Insulin resistance occurs when cells fail to respond effectively to insulin and is primarily associated with disruptions in insulin signaling pathways. In GDM, glucose uptake is impaired due to a reduction in GLUT4 translocation to the plasma membrane, with a reported 54% decrease compared to normal pregnancy (31). While insulin receptor expression remains intact, a reduction in tyrosine phosphorylation of the receptor disrupts insulin signaling and decreases insulin sensitivity (32). Additionally, alterations in the expression and phosphorylation of key insulin signaling regulators such as IRS-1, PI3K, and GLUT4 have been observed (14). Many of these molecular changes persist postpartum (33).

Neurohormonal dysfunction also plays a role in GDM pathogenesis. This complex network regulates appetite, energy expenditure, and basal metabolic rate through central (*e.g.*, cortical centers controlling cognitive, visual, and "reward" cues) and peripheral (*e.g.*, satiety and fasting hormones) signals. Dysregulation of this network contributes to GDM by affecting adiposity and glucose utilization. This system is largely governed by circadian rhythms, explaining the association between GDM and pathological sleep disorders or shift work. The most well-known hormones involved

are leptin and adiponectin. Leptin, secreted by adipocytes and the placenta, regulates energy homeostasis. Leptin levels are elevated in GDM, and hyperleptinemia has been associated with placental insulin resistance. Increased leptin levels have been reported to enhance fetal amino acid transport *via* the placenta, contributing to fetal macrosomia (34). Furthermore, leptin resistance has been reported in GDM pregnancies, impairing leptin signaling in peripheral tissues and negatively affecting glucose homeostasis (35). Conversely, adiponectin, a hormone that enhances insulin sensitivity and suppresses gluconeogenesis, is significantly reduced in GDM. This reduction is linked to β-cell dysfunction and insulin resistance (36). Moreover, decreased placental adiponectin expression has been suggested to contribute to fetal growth restriction, potentially counteracting the risk of macrosomia (37). Lower adiponectin levels correlate with maternal glucose intolerance, suggesting a pivotal role for this hormone in GDM pathogenesis (38).

Another key factor in GDM pathogenesis is altered adipose tissue storage. While early pregnancy is characterized by adipose tissue expansion, later gestation involves increased lipid mobilization to support fetal growth. These adipose tissue–related adaptations appear to become dysregulated in the context of GDM (39). The condition is characterized by impaired adipogenesis and exaggerated adipocyte hypertrophy, alongside suppression of critical components of the insulin signaling cascade, including transcription factors such as PPARγ, as well as fatty acid transport proteins.(40). The resulting decline in insulin sensitivity, coupled with defective adipocyte maturation, compromises the adipose tissue's capacity for efficient energy storage. This maladaptation promotes ectopic lipid deposition and exacerbates glucolipotoxicity in peripheral tissues, notably skeletal muscle and hepatic parenchyma, a feature shared with T2DM.

The placenta plays a crucial role in modulating insulin resistance during pregnancy by secreting hormones and cytokines (41). In GDM, the hyperglycemic intrauterine environment affects the transport of glucose, amino acids, and lipids across the placenta. Glucose transport across the placenta is insulin-independent, but maternal

hyperglycemia significantly influences this process, directly contributing to increased fetal growth and macrosomia (42). Similarly, alterations in system A and L amino acid transporters affect fetal growth (43). Additionally, most placental gene expression changes in GDM occur in lipid pathways, influencing glucose transport (44).

GDM has been reported to induce widespread hypermethylation in the placenta, triggering epigenetic modifications. However, the exact role of placental epigenetic and proteomic alterations remains unclear (45). Other factors implicated in GDM pathogenesis include gut microbiota, oxidative stress, and environmental toxin exposure.

BIOMARKERS AS PREDICTION TOOLS FOR DETERMINING GDM

Biomarkers are measurable biological substances that serve as indicators of normal physiological processes, pathological mechanisms, or pharmacodynamic responses to therapeutic interventions (46). They can be categorized as predictive markers, which evaluate the likelihood of disease development, or as diagnostic and screening tools that facilitate early detection during the subclinical phase, enabling timely intervention to prevent disease progression or anticipate treatment outcomes. Recent research has focused on biomarker-based approaches for improving the early detection of GDM. Various metabolic, inflammatory, and placental-derived biomarkers have shown promise in predicting GDM risk before hyperglycemia becomes clinically evident (47) (Figure 2).

Metabolic biomarkers

Given the established dysregulation of glucose metabolism in GDM, various metabolic markers have been explored to predict its onset. Elevated insulin resistance indices during the first trimester, assessed *via* the homeostasis model assessment of insulin resistance (HOMA-IR) using fasting serum glucose and insulin levels, have been linked to a heightened risk of developing GDM (48). A study demonstrated that increased HOMA-IR in early pregnancy is a significant risk factor for GDM, with specific cutoff

values varying based on body weight categories (49). However, insulin resistance alone may be insufficient as a predictive marker, as not all studies conclusively demonstrate this association.

Pregnancy naturally involves a progressive decline in insulin sensitivity, accompanied by increased insulin resistance, peaking in the third trimester and diminishing postpartum (32). Early gestational assessments of insulin sensitivity, employing indices such as the Matsuda index (derived from OGTT), quantitative insulin sensitivity check index, and HOMA for sensitivity, have been investigated as potential predictors of GDM (49). Given that alterations in insulin sensitivity precede GDM manifestation, further validation of these measures could facilitate early interventions before GDM develops (9).

Sex hormone-binding globulin (SHBG), a glycoprotein that binds androgens and estrogens, has been implicated in glucose metabolism and the development of type 2 diabetes (50). Low levels of SHBG are associated with insulin resistance and the development of GDM (51). owever, the predictive significance of SHBG diminishes when adjusting for factors such as body mass index (BMI), ethnicity, and family history, underscoring the necessity for biomarkers that offer predictive value beyond standard clinical risk factors (9).

Lipid metabolism undergoes significant changes during pregnancy, with the first and second trimesters marking substantial maternal fat accumulation due to increased lipid synthesis (52). Lipid levels rise progressively, peaking in late gestation (53). Triglycerides do not directly cross the placenta; however, placental lipoprotein receptors facilitate fatty acid transfer along the maternal-fetal gradient. In GDM, these alterations are more pronounced, with elevated triglyceride levels observed across all trimesters (54). Elevated maternal triglycerides in the third trimester are positively correlated with increased fetal birth weight, independent of GDM status (55). Conversely, higher maternal high-density lipoprotein (HDL) levels are inversely associated with the risk of fetal macrosomia (56). While the association between

disrupted maternal lipid profiles and GDM is established, their utility as predictive biomarkers requires further elucidation.

Given that GDM is a metabolic disorder, multiple studies have investigated metal ions, lipidomics, amino acids, metabolites, and vitamins as potential predictive biomarkers. Among metal ions evaluated in early pregnancy, iron and selenium demonstrated significantly different levels between GDM and normoglycemic pregnancies, with a sensitivity exceeding 80% (57). Elevated iron levels contribute to oxidative stress by increasing reactive oxygen species (58) production, whereas selenium, as a cofactor for antioxidant enzymes, supports placental function, and its deficiency may be linked to GDM development. The involvement of these redox-sensitive elements suggests a potential role for oxidative stress in GDM pathogenesis, likely associated with placental and mitochondrial dysfunction (57).

Nutritional status is crucial for maternal and fetal health. Lower vitamin D (25-hydroxyvitamin D) levels in the first trimester were associated with 81% sensitivity but only 44% specificity for GDM prediction, while elevated vitamin A levels in early pregnancy exhibited moderate predictive value (AUC = 0.649). Both vitamins A and D are implicated in immune regulation and insulin sensitivity, and their deficiency may contribute to GDM pathophysiology (59). Metabolomic profiling has identified several biomarkers for GDM prediction, with branched-chain amino acids (BCAAs)—valine, isoleucine, and leucine—demonstrating predictive values ≥ AUC 0.67 and a sensitivity exceeding 76%. Since BCAAs are linked to energy metabolism, their elevated levels correlate with insulin resistance and metabolic dysfunction, both central to GDM development (60).

Inflammatory biomarkers

Obesity is a well-established risk factor for GDM, characterized by a state of chronic low-grade inflammation due to excessive nutrient and energy intake (9). This pro-inflammatory state disrupts metabolic pathways in adipose tissue, the liver, and pancreas, altering levels of adipokines, chemokines, and cytokines (61, 62). Among

these, tumor necrosis factor-alpha (TNFα), secreted by the placenta, has been implicated in pregnancy-induced insulin resistance. A case-control study demonstrated that elevated maternal TNFα levels at 11–13 weeks of gestation were associated with increased GDM risk (63). Similarly, C-reactive protein (CRP), an inflammatory marker, has been investigated as a potential predictor of GDM, though its association weakens after BMI adjustment, limiting its specificity (64, 65). Interleukin-6 (IL-6), predominantly derived from adipocytes, is positively correlated with adiposity measures and insulin resistance (9). Even in the absence of maternal obesity, higher circulating IL-6 Levels have been linked to GDM development. However, its role as a prospective predictor of GDM remains uncertain, as studies have yet to establish whether its elevation precedes the onset of GDM or simply reflects its pathophysiology (66).

This large-scale retrospective study, involving 15,807 pregnant women, investigated the association between systemic inflammatory markers in early pregnancy and the risk of developing GDM. The findings revealed that women diagnosed with GDM exhibited lower monocyte counts and higher neutrophil and lymphocyte counts compared to the non-GDM group. Notably, monocyte count and monocyte-to-high-density lipoprotein ratio (MHR) were significantly associated with GDM risk, with a lower monocyte count and MHR correlating with an increased likelihood of GDM development. However, other inflammatory indices, including systematic immune inflammation index (SII) and systemic inflammation response index (SIRI), did not show a strong predictive value. Subgroup analysis further suggested that the association between monocyte count/MHR and GDM risk was particularly pronounced in women with a family history of diabetes. These findings propose that readily accessible inflammatory markers could serve as potential early indicators of GDM risk, highlighting the need for further validation in larger, multi-ethnic cohorts to establish their clinical utility in early GDM screening and prevention strategies (67).

Adipokine profiles in GDM pregnancies differ significantly from those in normoglycemic pregnancies, with decreased adiponectin and nesfatin-1 Levels, and increased leptin, resistin, visfatin, vaspin, and spexin (68). The predictive performance

of adipokines and cytokines varies widely, with AUC values ranging from 0.337 to 0.836, though 10 biomarkers exhibit an AUC > 0.7, indicating moderate predictive potential (68, 69). Among these, leptin and adiponectin are the most promising biomarkers, demonstrating 100% sensitivity in small cohort studies (70). Leptin, which regulates energy homeostasis and insulin sensitivity, is elevated in GDM, likely reflecting increased adiposity and insulin resistance. In contrast, adiponectin, which enhances insulin sensitivity, is significantly reduced, suggesting impaired glucose metabolism in GDM (71).

Placental Biomarkers

The placenta plays a pivotal role in inflammatory regulation, with its contribution becoming more pronounced in obesity, where it serves as a significant source of pro-inflammatory cytokines, including IL-1, TNFα, and IL-6 (9). Although the placenta has adaptive mechanisms to protect the fetus from inflammation , alterations in glucose transport have been observed, primarily mediated through modulation of glucose transporters (GLUTs) (72). In GDM pregnancies, placental GLUT9a expression is markedly increased, a phenomenon further exacerbated by exposure to exogenous insulin (73). Conversely, GLUT1 expression in the basal membrane remains stable within physiological glucose concentrations but is altered under extreme glycemic conditions (74). Despite these compensatory mechanisms, placental glucose uptake in GDM is elevated by 2–3-fold, reflecting a substantial adaptation to maternal metabolic disturbances (75). This evidence highlights the critical regulatory role of the placenta in modulating the fetal environment, attempting to mitigate the impact of maternal metabolic dysfunction at a cellular level. A deeper understanding of placental physiology may facilitate the identification of novel biomarkers for GDM and its associated fetal complications (9).

Placental dysfunction can disrupt the secretion of placental hormones, leading to an imbalance between pregnancy-induced insulin secretion and insulin resistance, thereby contributing to GDM development (76). Several placenta-derived proteins have shown

predictive potential for GDM (77). Pregnancy-associated plasma protein-A2 (PAPP-A2), measured at approximately 13.6 weeks of gestation, and placental protein 13 (PP13), assessed between 16–20 weeks, demonstrated sensitivities of 71% and 92%, respectively (78). PAPP-A2 regulates insulin-like growth factor activity, which is essential for fetal growth and glucose metabolism; its elevated levels may indicate placental stress or dysfunction, contributing to GDM onset (58). Similarly, PP13 is involved in placental development and vascularization, and altered levels may signal impaired placental function, predisposing to GDM (10). Emerging evidence supports the clinical utility of PAPP-A2 as a promising early biomarker for GDM. In a case-control study by Dereke *et al.*, circulating PAPP-A2 Levels were significantly elevated in women diagnosed with early-pregnancy GDM compared to matched normoglycemic controls [median 13.5 vs. 8.1 ng/mL; $P < 0.001$]. Importantly, PAPP-A2 Levels were independently associated with GDM after adjusting for age, BMI, C-peptide, and adiponectin. At a cut-off value of 6 ng/mL, the marker demonstrated a sensitivity of 96% and a negative predictive value of 99.7%, indicating strong potential for early risk stratification and reduction in unnecessary OGTTs. While these findings position PAPP-A2 as a compelling candidate for pre-screening, larger prospective cohorts are warranted to validate its performance across diverse populations (78).

Additionally, angiogenic factors, such as prokineticin 1 and soluble fms-like tyrosine kinase-1 (sFlt-1), have demonstrated strong predictive accuracy for GDM in the early second trimester, with sensitivities of 88% and 95%, respectively (10, 79). Prokineticin 1 plays a role in angiogenesis and placental function, and its dysregulation may impair placental blood flow and nutrient exchange, increasing GDM risk (79). sFlt-1, an antiangiogenic factor, disrupts placental vascularization, contributing to placental insufficiency and metabolic dysfunction (10). In contrast, placental growth factor (PlGF), a key biomarker for preeclampsia, has shown limited utility in predicting GDM, with a sensitivity of only 51%, suggesting it is not a reliable standalone predictor (80).

Afamin is a vitamin-E-binding glycoprotein, exhibited elevated concentrations in early pregnancy among GDM cases, reinforcing its link to oxidative stress and

metabolic dysfunction (81). Fatty acid-binding protein 4 (FABP4), galectin-3 (Gal-3), and fibronectin were also identified as potential markers, with higher FABP4 and Gal-3 Levels correlating with GDM onset, while fibronectin levels were lower in affected individuals (82-84). Furthermore, CD93 and HTRA-1, both implicated in angiogenesis and placental remodeling, showed altered expression in GDM, suggesting their involvement in disease pathogenesis (85, 86). Although these findings underscore the potential of biomarker-based GDM prediction, variability across studies, limited cohort sizes, and the influence of ethnic, environmental, and metabolic factors necessitate further large-scale, multi-ethnic validation studies to establish their clinical applicability.

Genetic and epigenetic biomarkers

Genetic predisposition plays a significant role in the pathogenesis of GDM, with numerous single nucleotide polymorphisms (SNPs) identified in genes related to insulin secretion, glucose metabolism, and adipose tissue function (87). The TCF7 L2 gene, a key regulator in the Wnt signaling pathway, has been strongly associated with both type T2DM and GDM. Among the most frequently studied polymorphisms, rs7903146 and rs4506565 have been linked to an increased risk of GDM, with individuals carrying the T allele of rs7903146 demonstrating higher proinsulin-to-insulin ratios, impaired insulinotropic effects of incretin hormones, and greater hepatic glucose production (88, 89). Similarly, MTNR1B, which encodes the melatonin receptor 1B, has been implicated in glucose intolerance and insulin resistance. The rs10830963 SNP in MTNR1B is frequently associated with elevated fasting glucose, increased HbA1c levels, and reduced early-phase insulin secretion, potentially modifying the effectiveness of lifestyle interventions in high-risk populations (90, 91). Additionally, ADIPOQ, a gene encoding adiponectin, has been linked to GDM risk through SNPs such as rs2241766, where the G allele is associated with reduced adiponectin levels and increased insulin resistance (92).

Beyond genetic predisposition, epigenetic modifications, particularly microRNAs (miRNAs), have emerged as crucial regulators in GDM development and progression. miRNAs are short non-coding RNAs that modulate gene expression post-transcriptionally, influencing β-cell function, insulin sensitivity, and inflammatory responses. Several miRNAs, including miR-29a, miR-29b, miR-132, and miR-223, have been differentially expressed in GDM pregnancies (93). Notably, miR-29a downregulation has been linked to impaired glucose metabolism through its role in regulating phosphoenolpyruvate carboxykinase 2, a key enzyme in gluconeogenesis (94). Moreover, miR-657 has been implicated in macrophage-mediated inflammation, promoting a pro-inflammatory environment that may contribute to GDM pathogenesis. Importantly, miRNA dysregulation appears to precede glucose abnormalities, suggesting their potential utility as early biomarkers for GDM diagnosis (95).

Further supporting the role of epigenetics, trimester-specific miRNA expression patterns have been observed in GDM pregnancies. Studies indicate that miR-517-3p and miR-518-5p exhibit increased levels in the second trimester, whereas their expression is downregulated in the third trimester, potentially reflecting dynamic regulatory mechanisms in placental function and maternal metabolic adaptation (96). Similarly, miR-125b-5p is upregulated in the first trimester but declines in the second trimester, highlighting the need for longitudinal studies to characterize miRNA fluctuations throughout pregnancy (97). The clinical utility of circulating microRNAs as predictive biomarkers for GDM has been supported by cohort-based validation studies. In a prospective analysis by Juchnicka *et al.*, serum samples from the first trimester were evaluated in 48 pregnant women, including 24 who later developed GDM and 24 normoglycemic controls (98). Using NanoString-based profiling followed by RT-PCR validation, the study identified three significantly upregulated microRNAs—miR-16-5p, miR-142-3p, and miR-144-3p—in the GDM group. ROC curve analysis demonstrated strong diagnostic performance, with AUC values of 0.868, 0.778, and 0.756, respectively. These findings reinforce the potential of microRNA panels for early, non-invasive GDM screening, and suggest that first-trimester expression signatures may precede clinical

glycemic alterations (98). Further supporting the utility of circulating miRNAs in early GDM prediction, Ye *et al*. identified a plasma exosomal miRNA panel—miR-122–5p, miR-148a-3p, miR-192–5p, and miR-99a-5p—that demonstrated robust discriminatory power in a nested case-control study conducted between 10 and 16 gestational weeks (AUC = 0.82). These miRNAs were significantly dysregulated in women who subsequently developed GDM and were mechanistically linked to insulin and AMPK signaling pathways. This study provides compelling evidence supporting the feasibility of non-invasive, early screening strategies based on extracellular vesicle-derived microRNAs (99).

These findings underscore the complex interplay between genetic susceptibility and epigenetic modifications in GDM, paving the way for biomarker-based risk stratification and personalized interventions. However, large-scale multi-ethnic cohort studies are required to validate these findings and establish standardized biomarker panels for clinical use.

Urine biomarkers

Recent advances in metabolomics and proteomics have facilitated the identification of potential urinary biomarkers for early prediction and diagnosis of GDM. Given its non-invasive collection and ability to reflect metabolic changes, urine has emerged as a promising biofluid for biomarker discovery (100). Studies have demonstrated significant alterations in amino acid metabolism, particularly BCAAs and tryptophan metabolites, in women who later develop GDM (101). Elevated urinary serotonin, kynurenine, and indole metabolites suggest dysregulation in the tryptophan-kynurenine pathway, which may contribute to IR and β-cell dysfunction (102). Similarly, increased purine metabolites such as hypoxanthine, xanthine, and uric acid have been linked to oxidative stress and chronic inflammation, key factors in GDM pathogenesis (103).Lipid metabolites, including ceramides, sphingomyelins, and glycerophospholipids, have also been identified as potential markers, reflecting dysregulated lipid metabolism and insulin signaling (104).

In addition to metabolomic changes, proteomic analyses have revealed several proteins with potential diagnostic value for GDM. Liver-type fatty acid-binding protein (L-FABP), a marker of renal stress and lipid metabolism, is significantly elevated in GDM patients, suggesting a role in early metabolic disturbances (105). Similarly, inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4), an acute-phase inflammatory protein, has been correlated with hyperglycemia severity and adverse neonatal outcomes (106). Moreover, coagulation factor IX and other fibrinolysis-related peptides were found to be upregulated in GDM, reflecting the pro-thrombotic state associated with hyperglycemia. These findings highlight the potential clinical utility of urinary biomarkers in identifying at-risk women before traditional diagnostic criteria are met, offering a window for early intervention (106).

Despite these promising discoveries, heterogeneity in findings, lack of large-scale validation, and inconsistencies due to diet, ethnicity, and analytical techniques remain key challenges. Future research should focus on standardized protocols for biomarker validation, longitudinal studies tracking urinary changes throughout pregnancy, and multi-omics integration to enhance predictive accuracy. Ultimately, the incorporation of urinary biomarkers into routine screening strategies could improve risk stratification, facilitate early therapeutic interventions, and mitigate the long-term metabolic consequences of GDM (100).

MACHINE LEARNING-BASED PREDICTIONS FOR GDM

Machine learning (ML) has emerged as a promising approach for predicting GDM, offering an alternative to traditional screening methods. Recent meta-analysis evaluated 25 studies that developed and validated ML-based models for GDM prediction, demonstrating an overall pooled area under the receiver operating characteristic curve of 0.8492, indicating strong discriminatory power. Among the various ML techniques analyzed, non-logistic regression models, including gradient-boosting decision trees (GBDT), support vector machines, and artificial neural networks, outperformed traditional logistic regression models, achieving an AUROC of 0.8891 compared to

0.8151 for logistic regression. Key predictors consistently identified across studies included maternal age, BMI, family history of diabetes, and fasting blood glucose levels, with additional emerging predictors such as triglycerides, PAPP-A, and leptin levels (107).

The analysis highlighted that early pregnancy ML models (0-13 weeks gestation) achieved the highest sensitivity (0.74), while models trained on data from 14-28 weeks gestation exhibited greater specificity (0.85), suggesting that ML-based tools may facilitate both early risk stratification and mid-gestation confirmatory screening. Notably, ensemble methods such as LightGBM and GA-CatBoost yielded superior predictive performance compared to decision trees (DT) and k-nearest neighbors, which struggled with high-dimensional clinical datasets. The study further emphasized that feature selection and data preprocessing are critical in optimizing ML model performance. While external validation was performed in only 16% of included studies, findings underscore the need for standardized diagnostic criteria, prospective validation, and real-world implementation studies to enhance clinical applicability (107)
.

Despite the demonstrated potential of ML in improving GDM prediction, challenges remain. Heterogeneity among studies, differences in population characteristics, and inconsistencies in diagnostic criteria limit the generalizability of findings. Furthermore, the integration of ML models into clinical workflows necessitates robust validation, regulatory approval, and physician interpretability. Future research should focus on incorporating multi-omics data, wearable technology-derived biomarkers, and explainable AI approaches to further refine risk prediction and support personalized prenatal care strategies (107).

Machine learning-based predictions for GDM progression to T2DM

The application of ML algorithms in predicting the progression from GDM to T2DM has gained increasing attention due to its potential for early risk stratification and personalized intervention (108). Recent meta-analysis evaluated 13 studies involving

11,320 women with a history of GDM, assessing the predictive accuracy of 22 ML models. The findings demonstrated that ML models achieved a pooled C-statistic of 0.82 (95%CI: 0.79–0.86), indicating strong discriminative ability for identifying women at high risk of developing T2DM. The pooled sensitivity (76%) was considerably higher than traditional risk assessment tools, though specificity remained moderate at 57%, highlighting the need for refinement in predictive models (108).

Recent research has demonstrated the growing potential of machine learning ML models to identify women at high risk of GDM using early pregnancy data. Zaky *et al*. developed a stacked ensemble model incorporating 26 variables, including HOMA-IR, insulin, and NT-proBNP levels, from the Qatar Birth Cohort. Their model achieved a recall of 92.1% and an accuracy of 88.8%, illustrating the feasibility of early, non-invasive GDM prediction (109). Similarly, Kaya *et al*. retrospectively analyzed data from a Turkish tertiary center, using maternal sociodemographic and obstetric variables collected during the first trimester. Among multiple algorithms tested, the XGBoost classifier demonstrated the best performance, with accuracies of 66.7% and 72.7% and AUCs of 55% and 73.3% in nulliparous and primiparous subgroups, respectively (110). Both studies highlighted key predictive features such as fasting plasma glucose, BMI, and family history of diabetes. Despite promising results, challenges such as limited sample size, lack of external validation, and variability in electronic health records remain barriers to clinical translation. These examples underscore the need for standardized data pipelines, population-specific models, and prospective multicenter validation before ML tools can be effectively integrated into GDM screening protocols.

Different ML algorithms, including logistic regression, DT, random forest, Naïve Bayes, and Cox models, were compared. Among these, random forest exhibited the highest predictive accuracy (C-statistic: 0.88, 95%CI: 0.81–0.94), whereas logistic regression models showed slightly lower performance (C-statistic: 0.81, 95%CI: 0.75–0.88). Key predictive variables incorporated in these models included maternal metabolic markers, glucose and lipid metabolites, genetic factors, and traditional clinical risk factors. Notably, the inclusion of circulating microRNAs (miRNAs] such as

miR-369-3p and miR-543 significantly enhanced model performance by improving early risk assessment (108).

Despite promising results, several challenges remain in the clinical translation of ML models. The heterogeneity of study designs, variability in follow-up durations, and inconsistent validation protocols introduce potential bias and limit generalizability. Moreover, the lack of external validation in most models raises concerns regarding their real-world applicability across diverse populations (108).Future research should prioritize large-scale prospective studies, integration of multi-omics data, and development of standardized ML algorithms to refine risk prediction and enhance clinical decision-making for postpartum women with a history of GDM. Ultimately, the incorporation of ML-based predictive models into clinical workflows could facilitate targeted interventions, reduce the burden of T2DM, and improve long-term maternal metabolic health outcomes.

CONCLUSION

Extensive research into the molecular pathogenesis of GDM has led to the identification of a broad spectrum of biomarkers with potential utility for early prediction and diagnosis. These include metabolic indices such as HOMA-IR and SHBG, lipid profiles, adipokines like leptin and adiponectin, placental proteins including PAPP-A2 and PP13, angiogenic factors such as sFlt-1 and prokineticin 1, as well as a range of microRNAs and urinary metabolites. Many of these markers exhibit significant correlations with insulin resistance, β-cell dysfunction, or placental insufficiency, all of which contribute to GDM development. Although several candidates have shown moderate-to-high sensitivity and specificity in small- to medium-scale studies, clinical translation remains limited due to heterogeneity in study populations, biomarker variability across gestational ages, and lack of standardized thresholds. Nonetheless, the incorporation of validated biomarker panels into existing screening protocols may improve early risk stratification and allow for timely intervention prior to the clinical onset of hyperglycemia.

Simultaneously, ML and artificial intelligence (AI) models are transforming the landscape of GDM risk prediction and diagnosis . Traditional risk assessment tools rely on linear models and clinical risk factors, but ML algorithms can process high-dimensional. Despite these advancements, several challenges remain in translating biomarker discoveries and AI-driven risk assessment tools into clinical practice. Issues such as data standardization, external validation, regulatory approvals, and healthcare integration need to be addressed to ensure that these methods are cost-effective, accessible, and widely applicable across diverse populations. Future research should focus on multi-cohort validation studies, AI integration with wearable health technologies, and personalized follow-up protocols to enhance GDM screening and management strategies.

While several recent reviews have examined aspects of GDM prediction and screening, they have either narrowly focused on traditional risk models or discussed methodological limitations of emerging tools. For example, Germaine *et al*. emphasized challenges in implementing ML models—such as lack of data transparency and external validation—but did not address biomarker-based prediction strategies or their integration with ML frameworks (111). Similarly, ADA reviewed screening thresholds and clinical criteria but included only limited discussion of conventional biomarkers like CRP and HbA1c, without exploring recent advancements in placental, metabolic, genetic, or microRNA-based diagnostics (5). In contrast, our review provides a comprehensive synthesis of validated and emerging biomarkers, while also evaluating state-of-the-art ML applications, thus offering a translational perspective that bridges biological discovery and computational modeling for early, individualized GDM risk stratification.

111309_Revision_Auto_Edited.docx

ORIGINALITY REPORT

9%
SIMILARITY INDEX

PRIMARY SOURCES

1	Hasini Rathnayake, Luhao Han, Fabrício da Silva Costa, Cristiane Paganoti et al. "Advancement in predictive biomarkers for gestational diabetes mellitus diagnosis and related outcomes: a scoping review", BMJ Open, 2024 Crossref	116 words — 2%
2	www.ncbi.nlm.nih.gov Internet	108 words — 2%
3	pmc.ncbi.nlm.nih.gov Internet	47 words — 1%
4	journals.lww.com Internet	32 words — 1%
5	www.livescience.com Internet	20 words — < 1%
6	doaj.org Internet	19 words — < 1%
7	www.racgp.org.au Internet	16 words — < 1%
8	J. Dereke, C. Nilsson, H. Strevens, M. Landin-Olsson, M. Hillman. "Pregnancy-associated plasma protein-A2 levels are increased in early-pregnancy gestational diabetes: a novel biomarker for early risk estimation", Diabetic Medicine, 2019 Crossref	15 words — < 1%

9	clinical-laboratory-diagnostics.com Internet	14 words — < 1%
10	coek.info Internet	14 words — < 1%
11	oaktrust.library.tamu.edu Internet	14 words — < 1%
12	www.science.gov Internet	14 words — < 1%
13	academic.oup.com Internet	13 words — < 1%
14	link.springer.com Internet	13 words — < 1%
15	www.cureus.com Internet	13 words — < 1%
16	www.hcplive.com Internet	13 words — < 1%
17	www.slideshare.net Internet	13 words — < 1%
18	TARUNA IKRAR. "A Double-Point Mutation in the Selectivity Filter Site of the KCNQ1 Potassium Channel Results in a Severe Phenotype, LQT1, of Long QT Syndrome", Journal of Cardiovascular Electrophysiology, 2/5/2008 Crossref	12 words — < 1%
19	acikerisim.istanbul.edu.tr:8080 Internet	12 words — < 1%
20	ccforum.biomedcentral.com Internet	12 words — < 1%

21	lup.lub.lu.se Internet	12 words — < 1%
22	www.frontiersin.org Internet	12 words — < 1%
23	www.nature.com Internet	12 words — < 1%
24	www.nepho.org Internet	12 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 12 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS