07/07/2025

We would like to sincerely thank the reviewers for their thoughtful and constructive comments. We have carefully addressed each of the points raised and revised the manuscript accordingly. A detailed, point-by-point response is provided below.

1- Enhance Machine Learning Section: The section on machine learning-based predictions for GDM and its progression to type 2 diabetes mellitus (T2DM) is promising but could be strengthened by including specific examples or case studies of successful clinical implementations. For instance, discussing a study where machine learning improved diagnostic accuracy in a real-world setting would add practical context. Additionally, elaborating on challenges, such as data standardization or model validation, would provide a more balanced perspective. Consider citing recent studies, such as those available on PubMed, to support this section.

Respond: We thank the reviewer for this valuable suggestion. In response, we have expanded the machine learning section to include two recent studies—Zaky *et al*. (2025) and Kaya *et al*. (2024)—which illustrate real-world applications of ML-based models for early GDM prediction. These examples provide practical context and highlight both the diagnostic potential and the implementation challenges of ML approaches, including data standardization and external validation. The revised content has been highlighted in the manuscript.

2- Deepen Biomarker Validation Discussion: The biomarker section categorizes metabolic, inflammatory, placental, genetic, and urinary markers effectively. However, including more details on their validation—such as performance in large-scale cohort studies or ongoing clinical trials—would enhance credibility. For example, discussing the sensitivity and specificity of biomarkers like PAPP-A2 or microRNAs in diverse populations could clarify their clinical utility. Resources like ClinicalTrials.gov may provide relevant trial data to reference.

Respond: In response, we have expanded the microRNA section by incorporating two cohort-based studies—Juchnicka *et al*. and Ye *et al*.—which provide validation data,

including AUC values and mechanistic associations with GDM pathophysiology. The revised content has been highlighted in the manuscript.

3- Expand Future Directions: The conclusion mentions integrated, multi-omic approaches, which is forward-thinking. Expanding this section with specific research suggestions—such as interdisciplinary collaborations between bioinformaticians and clinicians or studies focusing on underrepresented populations—would strengthen its impact. For example, proposing research on machine learning models tailored to high-risk ethnic groups could address current gaps.

Respond: In the revised conclusion, we have highlighted the limitations of current screening strategies recommended by the ADA and proposed future directions focused on biomarker-based and machine learning–driven approaches. These additions underscore the potential for shifting from conventional diagnostic thresholds to individualized risk stratification models. The updated content has been clearly marked in the manuscript.

4- Verify Reference Completeness: The reference list is assumed to be comprehensive, but ensure it includes the most recent studies, particularly in fast-evolving areas like biomarkers and machine learning. A quick check for seminal papers from 2023–2025, accessible *via* Google Scholar, could confirm this. Avoid over-reliance on older references or self-citations, ensuring a balanced representation of the field.

Respond: In the course of our revisions, we have already incorporated several recent studies published between 2022 and 2025, particularly in the domains of biomarkers and machine learning. We have carefully reviewed the reference list to ensure it reflects current and balanced representation of the field.

5- Image and Box Quality: The figures and boxes are described as clear and necessary, with consecutive numbering (*e.g.*, Figure 1, Figure 2; Box 1–4). However, please confirm that image resolution meets journal standards (*e.g.*, 300 dpi for print). Additionally, ensure no redundancy exists between text and visual content, as this could streamline the manuscript.

Respond: We appreciate the reviewer's careful evaluation. We have reviewed all figures and ensured that the resolution meets the required 300 dpi standard. We also revised the text to minimize redundancy with visual elements and confirm that figures and boxes are consecutively numbered and integrated appropriately.

Round 2

*Thank you for submitting your revised manuscript, "Advances in gestational diabetes mellitus screening: Emerging trends and future directions," to the *World Journal of Diabetes*. Your thorough responses to the reviewers' comments and the subsequent revisions have significantly strengthened the manuscript. Below are specific comments addressing the revised content, the quality of the revisions, and suggestions for further refinement to ensure the manuscript meets the journal's standards for publication.*

1. Response to Reviewer Comment 1: Enhancing the Machine Learning Section

Your inclusion of Zaky et al. (2025) and Kaya et al. (2024) as examples of real-world applications of machine learning (ML) for early GDM prediction is commendable. These additions effectively illustrate the diagnostic potential of ML models and address implementation challenges, such as data standardization and external validation. However, to further enhance this section, consider briefly discussing the specific ML algorithms used in these studies (e.g., random forests, neural networks) and their performance metrics (e.g., AUC, sensitivity, specificity) to provide readers with a clearer understanding of their practical utility. Additionally, a brief mention of how these models could be integrated into clinical workflows would strengthen the practical context. For example, you could note whether these models are deployable in resource-limited settings or require advanced computational infrastructure.

2. Response to Reviewer Comment 2: Deepening Biomarker Validation Discussion

The expanded discussion on microRNA biomarkers, incorporating Juchnicka et al. (2022) and Ye et al. (2022), is a valuable addition. The inclusion of AUC values and mechanistic associations with GDM pathophysiology enhances the credibility of the biomarker section. However, to further improve this section, consider adding a brief comparison of microRNA biomarkers with other biomarker categories (e.g., metabolic or inflammatory markers) to clarify their relative predictive power. For instance, are microRNAs more sensitive or specific than HOMA-IR or inflammatory markers like CRP? Additionally, referencing ongoing clinical trials from ClinicalTrials.gov, as suggested by the reviewer, could further validate the clinical utility of these

biomarkers. If no relevant trials are available, a statement acknowledging this gap would demonstrate thoroughness.

3. Response to Reviewer Comment 3: Expanding Future Directions

The revised conclusion effectively highlights limitations in current ADA-recommended screening strategies and proposes biomarker-based and ML-driven approaches for individualized risk stratification. The focus on shifting from conventional diagnostic thresholds to personalized models is forward-thinking and aligns well with emerging trends in precision medicine. To further strengthen this section, consider specifying one or two high-risk ethnic groups (e.g., South Asian or Hispanic populations) and discussing how ML models could be tailored to address their unique risk profiles. Additionally, a brief mention of potential barriers to implementing multi-omic approaches (e.g., cost, accessibility, or regulatory challenges) would provide a more balanced perspective.

4. Response to Reviewer Comment 4: Verifying Reference Completeness

The updated reference list, incorporating studies from 2022–2025, demonstrates a robust effort to reflect current research in biomarkers and ML. The inclusion of recent studies such as Zhao et al. (2025) and Zaky et al. (2025) strengthens the manuscript's relevance. However, to ensure a truly balanced representation, consider checking for any seminal papers from underrepresented research groups or regions (e.g., Africa or South America) to address global perspectives on GDM. A quick Google Scholar search for 2023–2025 studies from these regions could confirm whether additional references are needed. Additionally, ensure that self-citations, if present, are minimal and justified to avoid potential bias.

5. Response to Reviewer Comment 5: Image and Box Quality

Your confirmation that all figures meet the journal's 300 dpi resolution standard and that redundancy between text and visual content has been minimized is appreciated. The consecutive numbering of figures (e.g., Figure 1, Figure 2) and boxes (e.g., Box 1–4) is clear and appropriate. However, based on the provided PowerPoint file ("111309-Image File-revision.pptx"), the content on Page 2 lists biomarker categories (e.g.,

metabolic, inflammatory, placental, genetic, urinary) but contains a duplication of "Genetic Biomarker S" and "Urine Biomarker S." This appears to be an error and should be corrected to avoid confusion. Additionally, consider adding a brief caption or legend to Figure 2 (Biomarkers in GDM) in the manuscript to clarify the abbreviations (e.g., SII, SIRI, FABP4) for readers unfamiliar with these terms. If not already included, ensure that all figures are referenced in the text to guide readers effectively.

6. General Comments on Manuscript Quality

The manuscript is well-structured, with a clear introduction, detailed pathogenesis section, and comprehensive discussion of biomarkers and ML-based approaches. The core tip and abstract succinctly summarize the manuscript's contributions, emphasizing precision diagnostics and individualized care. The scientific quality and significance are high (Grade A), as the manuscript synthesizes current evidence and identifies critical gaps in GDM screening. However, the novelty and creativity (Grade B) could be further enhanced by explicitly addressing how the proposed approaches differ from existing reviews or by proposing a novel framework for integrating biomarkers and ML in clinical practice.

7. Suggestions for Further Improvement

Clarify Ethnic Variability: Given the global variability in GDM prevalence (6–20%), consider adding a brief subsection or paragraph discussing how biomarkers and ML models perform across diverse populations (e.g., Asian vs Caucasian vs African cohorts). This would address the reviewer's suggestion to focus on underrepresented populations and enhance the manuscript's global relevance.

Streamline Biomarker Section: The biomarker section is comprehensive but lengthy. Consider consolidating overlapping content (e.g., metabolic and inflammatory markers) into a table summarizing key biomarkers, their predictive performance (e.g., sensitivity, specificity), and validation status. This would improve readability and allow readers to quickly compare biomarker utility.

Visual Aids: If feasible, include a simple chart to visualize the performance of ML models (e.g., AUC values from Zaky et al. and Kaya et al.) or biomarker sensitivity/specificity across studies. This would enhance the manuscript's visual appeal and support the reviewer's emphasis on practical context.

Example Chart for ML Model Performance (if desired):

If you would like to include a chart comparing the performance of ML models from Zaky et al. (2025) and Kaya et al. (2024), I can generate one. For example, a bar chart could display AUC values for each model. Please confirm if you would like me to produce this chart, and provide specific AUC values if available (or I can note that these should be extracted from the cited studies).

8. Formatting and Minor Edits
Ensure consistency in terminology (e.g., "GDM" vs "gestational diabetes mellitus") throughout the manuscript to avoid redundancy. For instance, after defining GDM, use the acronym consistently unless introducing a new section.

In the abstract, consider replacing "multi-dimensional biomarkers" with "multimodal biomarkers" for clarity and alignment with scientific terminology.

*Double-check the reference format to ensure compliance with the *World Journal of Diabetes* guidelines (e.g., correct PMID and DOI formatting).*

Conclusion:
*The revised manuscript is of high scientific quality and addresses the reviewers' comments effectively. The additions to the machine learning and biomarker sections, along with the expanded future directions, significantly enhance the manuscript's impact. Addressing the minor issues noted above (e.g., correcting the duplicated biomarker categories, adding captions for figures, and considering ethnic variability) will further polish the manuscript for publication. I recommend **acceptance with*

minor revisions**. Please clarify whether you would like a chart generated for ML model performance or any other specific additions.

Thank you for your diligent work on this important topic. I look forward to seeing the final version of your manuscript.

12/08/2025 We would like to sincerely thank the reviewers for their thoughtful and constructive comments. We have carefully addressed each of the points raised and revised the manuscript accordingly. A detailed, point-by-point response is provided below.

1- Your inclusion of Zaky *et al*. (2025) and Kaya *et al*. (2024) as examples of real-world applications of machine learning (ML) for early GDM prediction is commendable. These additions effectively illustrate the diagnostic potential of ML models and address implementation challenges, such as data standardization and external validation. However, to further enhance this section, consider briefly discussing the specific ML algorithms used in these studies (*e.g.*, random forests, neural networks) and their performance metrics (*e.g.*, AUC, sensitivity, specificity) to provide readers with a clearer understanding of their practical utility. Additionally, a brief mention of how these models could be integrated into clinical workflows would strengthen the practical context. For example, you could note whether these models are deployable in resource-limited settings or require advanced computational infrastructure.

Respond: Thank you for your constructive suggestion. Performance metrics such as accuracy, recall, and AUC were already included in the original revision for both Zaky *et al*. and Kaya *et al*. To further strengthen this section, we have now added a brief note on the infrastructural requirements and clinical applicability of ML models, particularly in low-resource settings. These changes aim to clarify the practical context of integrating such tools into real-world GDM screening workflows.

2- The expanded discussion on microRNA biomarkers, incorporating Juchnicka *et al*. (2022) and Ye *et al*. (2022), is a valuable addition. The inclusion of AUC values and mechanistic associations with GDM pathophysiology enhances the credibility of the biomarker section. However, to further improve this section, consider adding a brief comparison of microRNA biomarkers with other biomarker categories (*e.g.*, metabolic or inflammatory markers) to clarify their relative predictive power. For instance, are microRNAs more sensitive or specific than HOMA-IR or inflammatory markers like CRP? Additionally, referencing ongoing clinical trials from ClinicalTrials.gov, as suggested by the reviewer, could further validate the clinical utility of these biomarkers. If no relevant trials are available, a statement acknowledging this gap would demonstrate thoroughness.

Respond: We have added a brief comparison paragraph at the end of the microRNA biomarker subsection, summarizing the relative predictive performance of microRNAs compared to other biomarker categories discussed in our review. This comparison was based entirely on the data already presented in the manuscript, highlighting that microRNAs generally show higher AUC values and more consistent predictive potential than many single metabolic or inflammatory markers. Additionally, we conducted a literature search for clinical trials directly comparing microRNA biomarkers with other biomarker types in GDM prediction; however, no such studies were identified.

3- The revised conclusion effectively highlights limitations in current ADA-recommended screening strategies and proposes biomarker-based and ML-driven approaches for individualized risk stratification. The focus on shifting from conventional diagnostic thresholds to personalized models is forward-thinking and aligns well with emerging trends in precision medicine. To further strengthen this section, consider specifying one or two high-risk ethnic groups (*e.g.*, South Asian or Hispanic populations) and discussing how ML models could be tailored to address their unique risk profiles. Additionally, a brief mention of potential barriers to implementing multi-omic approaches (*e.g.*, cost, accessibility, or regulatory challenges) would provide a more balanced perspective.

Respond: In the revised conclusion, we have included South Asian populations as an example of a high-risk ethnic group and discussed how ML-based models could be tailored to address their specific risk profile. Additionally, we have added a brief note on potential barriers to implementing multi-omic strategies, such as high cost, limited accessibility, and lack of regulatory harmonization, to provide a more balanced perspective.

4- Verify Reference Completeness: The updated reference list, incorporating studies from 2022–2025, demonstrates a robust effort to reflect current research in biomarkers and ML. The inclusion of recent studies such as Zhao *et al*. (2025) and Zaky *et al*. (2025) strengthens the manuscript's relevance. However, to ensure a truly balanced representation, consider checking for any seminal papers from underrepresented research groups or regions (*e.g.*, Africa or South America) to address global perspectives on GDM. A quick Google Scholar search for 2023–2025 studies from these regions could confirm whether additional references are needed. Additionally, ensure that self-citations, if present, are minimal and justified to avoid potential bias.
Respond: In line with the recommendation, we have expanded the global perspective by including recent literature addressing high-risk South Asian populations, which are notably underrepresented in many GDM studies. While the current scope of our review primarily focuses on regions with the highest volume of biomarker and ML research, we ensured that the updated reference list reflects diverse geographic contributions where relevant to the manuscript's objectives. Self-citations were kept to a minimum and included only when directly pertinent to the discussed concepts.

5- Your confirmation that all figures meet the journal's 300 dpi resolution standard and that redundancy between text and visual content has been minimized is appreciated. The consecutive numbering of figures (*e.g.*, Figure 1, Figure 2) and boxes (*e.g.*, Box 1–4) is clear and appropriate. However, based on the provided PowerPoint file ("111309-Image File-revision.pptx"), the content on Page 2 lists biomarker categories (*e.g.*, metabolic, inflammatory, placental, genetic, urinary) but contains a duplication of "Genetic Biomarker S" and "Urine Biomarker S." This appears to be an

error and should be corrected to avoid confusion. Additionally, consider adding a brief caption or legend to Figure 2 (Biomarkers in GDM) in the manuscript to clarify the abbreviations (*e.g.*, SII, SIRI, FABP4) for readers unfamiliar with these terms. If not already included, ensure that all figures are referenced in the text to guide readers effectively.

Respond: Upon re-examination of Figure 2, we confirm that FAB4 is listed under Urine Biomarkers in the current version. However, given that FABP4 (Fatty Acid Binding Protein 4) has been studied in both urinary and placental contexts, its categorization could arguably fit either group. For clarity, the full term corresponding to each abbreviation, including FABP4, has been provided in the PowerPoint slide notes submitted with the revision, allowing the editorial team to review and verify all biomarker definitions.

6- The manuscript is well-structured, with a clear introduction, detailed pathogenesis section, and comprehensive discussion of biomarkers and ML-based approaches. The core tip and abstract succinctly summarize the manuscript's contributions, emphasizing precision diagnostics and individualized care. The scientific quality and significance are high (Grade A), as the manuscript synthesizes current evidence and identifies critical gaps in GDM screening. However, the novelty and creativity (Grade B) could be further enhanced by explicitly addressing how the proposed approaches differ from existing reviews or by proposing a novel framework for integrating biomarkers and ML in clinical practice.

Respond: We have explicitly addressed how our proposed strategies differ from existing reviews by incorporating a novel, stepwise integration framework. This framework positions validated biomarker panels as primary triage tools, followed by ethnicity-specific ML risk models dynamically updated with longitudinal clinical and wearable-derived data, thereby providing a translational approach that bridges biological discovery and computational modeling.

7- Clarify Ethnic Variability: Given the global variability in GDM prevalence (6–20%), consider adding a brief subsection or paragraph discussing how biomarkers and ML

models perform across diverse populations (*e.g.*, Asian *vs* Caucasian *vs* African cohorts). This would address the reviewer's suggestion to focus on underrepresented populations and enhance the manuscript's global relevance. Streamline Biomarker Section: The biomarker section is comprehensive but lengthy. Consider consolidating overlapping content (*e.g.*, metabolic and inflammatory markers) into a table summarizing key biomarkers, their predictive performance (*e.g.*, sensitivity, specificity), and validation status. This would improve readability and allow readers to quickly compare biomarker utility. Visual Aids: If feasible, include a simple chart to visualize the performance of ML models (*e.g.*, AUC values from Zaky *et al*. and Kaya *et al*.) or biomarker sensitivity/specificity across studies. This would enhance the manuscript's visual appeal and support the reviewer's emphasis on practical context.
Respond: We have incorporated additional content addressing ethnic variability in GDM prevalence and biomarker/ML model performance, with specific reference to South Asian and other high-risk populations, to enhance the manuscript's global relevance. As suggested, we have prepared and incorporated a bar chart comparing the AUC values reported by Zaky *et al*. (2025) and Kaya *et al*. (2024) to visually illustrate the performance of different ML models for GDM prediction.

8- Ensure consistency in terminology (*e.g.*, "GDM" *vs* "gestational diabetes mellitus") throughout the manuscript to avoid redundancy. For instance, after defining GDM, use the acronym consistently unless introducing a new section. In the abstract, consider replacing "multi-dimensional biomarkers" with "multimodal biomarkers" for clarity and alignment with scientific terminology. Double-check the reference format to ensure compliance with the **World Journal of Diabetes** guidelines (*e.g.*, correct PMID and DOI formatting)
Respond: The term "gestational diabetes mellitus" (GDM) is now defined at its first occurrence in the manuscript, including at the beginning of the relevant paragraph, and the acronym "GDM" is used consistently thereafter to avoid redundancy. "multi-dimensional biomarkers" has been replaced with the scientifically more precise term "multimodal biomarkers."

Round 2

#08415436

Revision review report: 1. Nutritional interventions are not adequately addressed in the article. More emphasis on the preventive effects of simple sugar consumption and dietary and lifestyle interventions (e.g., low-glycemic-index diet, Mediterranean diet, exercise) could have enriched the article. I believe adding a few paragraphs or headings would also increase your citation count.

Respond: We sincerely appreciate the reviewer's insightful suggestion regarding the inclusion of nutritional and lifestyle interventions. We fully agree that dietary factors, such as simple sugar consumption, low-glycemic-index or Mediterranean diets, and physical activity, play an important role in the prevention and management of GDM. However, the scope of our review was specifically defined to address screening and diagnostic strategies, with a particular focus on conventional OGTT-based approaches, novel biomarkers, and machine learning–based prediction models.

Expanding the manuscript to include preventive and therapeutic interventions would necessitate a comprehensive discussion of nutritional, lifestyle, and pharmacological management strategies. This would considerably broaden the scope and length of the article, potentially diverting the focus away from its primary aim of evaluating emerging trends in GDM screening. We therefore respectfully suggest that these important aspects are best addressed in a separate review dedicated to prevention and treatment modalities, whereas the current manuscript remains focused on diagnostic and predictive frameworks.

Revision review report: 2. Numerous AUC values, logistic regression results, and statistical data are presented. These details are excessive in a review format, distracting the reader, and giving the text a thesis-like feel. (for example, this section: 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 (DT), support vector machines, and artificial neural networks, outperformed traditional logistic regression models, achieving an area under the receiver operating characteristic of 0.8891 compared to 0.8151 for logistic regression.

Other examples: BCAAs demonstrated predictive performance of AUC ≥ 0.67 with sensitivity above 76%, adipokines varied widely with AUCs between 0.337 and 0.836, and vitamin D deficiency showed a sensitivity of 81% but low specificity (44%)[60]. 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 (χ2: 0.88, 95%CI: 0.81-0.94), whereas logistic regression models showed slightly lower performance (χ2: 0.81, 95%CI: 0.75-0.88). these explanations are too long and distracting)

Respond: We thank the reviewer for this valuable comment. Following your suggestion, we simplified overly detailed statistical descriptions (e.g., AUC values, χ^2 results) to more concise narrative forms, while retaining only key representative numbers for clarity. These revisions, highlighted in yellow in the manuscript, aim to improve readability and align the text with the review format. We hope that these revisions satisfactorily address the remaining concerns and further enhance the quality of the manuscript.

We greatly appreciate your guidance and consideration.

JOURNAL EDITORIAL BOARD'S REVIEW REPORT

Dear Editor, Journal edıtorıal board comments to authors:

The comments provided by the two reviewers are thoughtful and perceptive.

The revised manuscript has adequately addressed the majority of the reviewers' concerns. This manuscript holds broad relevance to the medical field and presents general significance. The topic, which focuses on advancements in gestational diabetes mellitus (GDM) screening—incorporating machine learning (ML) approaches and biomarkers for GDM prediction—demonstrates high scientific quality and novelty. One point requires further reconsideration. A reviewer advised the authors to discuss the performance of biomarkers and ML models across heterogeneous populations (e.g., Asian, Caucasian, and African cohorts). Nevertheless, the authors' revision remains insufficient. Although they incorporated a single study investigating the prediction of postpartum diabetes in South Asian women with GDM utilizing ML methodologies, they did not furnish a comparative analysis of ML model performance across distinct populations. This needs further revision. So, I think a minor revision should be accomplished to purify this manuscript.

Respond: As suggested, we expanded the machine learning section by adding a new paragraph discussing the performance of biomarkers and ML models across heterogeneous populations. Additionally, to improve the flow of the manuscript, one paragraph in the discussion was repositioned, which led to changes in the reference order. These modifications are highlighted in blue within the revised text. In addition, we have added the full forms of abbreviations to Figure 1 for clarity.

With kind regards,

A.Seval Ozgu-Erdinc, Prof. Dr.

University Of Health Sciences, Ankara Bilkent City Hospital

Perinatology Department
ORCID iD:0000-0002-6132-5779