Received: 22 May 2025 | Accepted: 11 September 2025

DOI: 10.1111/dme.70173

DIABETIC Medicine

SYSTEMATIC REVIEW

Metformin safety during pregnancy in women with gestational diabetes mellitus: A systematic review and meta-analysis of maternal, neonatal and long-term outcomes

Mathilde Louise Saxdorff Brinkmann[1,2] | **Nicoline Josefine Krasilnikoff**[1,2] | **Mariam Ali Chaaban**[3] | **Birgitte Moeller Luef**[1,2] | **Christine Newman**[4,5] | **Fidelma Dunne**[4,5] | **Dorte Moeller Jensen**[2,3] | **Christina Anne Vinter**[1,2,3]

[1]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[2]Department of Gynecology and Obstetrics, Odense University Hospital, Odense, Denmark

[3]Steno Diabetes Center Odense, Odense University Hospital, Odense, Denmark

[4]HRB Clinical Research Facility Galway, Galway University Hospital, Galway, Ireland

[5]College of Medicine, Nursing and Health Sciences, University of Galway, Galway, Ireland

Correspondence
Nicoline Josefine Krasilnikoff, Department of Clinical Research, University of Southern Denmark, Odense, Denmark.
Email: nicolinekrasilnikoff@hotmail.com

Abstract

Aims: This systematic review and meta-analysis assessed the safety and efficacy of metformin in managing gestational diabetes mellitus (GDM), focusing on maternal, neonatal and long-term outcomes. While lifestyle changes are first-line treatment, pharmacological therapy is often required. Insulin, the standard, has drawbacks including weight gain, neonatal hypoglycaemia and maternal anxiety. Metformin is a promising alternative due to its insulin-sensitizing effects, but concerns remain about placental transfer and long-term effects on offspring.

Methods: A systematic search was conducted in PubMed and Embase up to 29 August 2024, including randomized controlled trials (RCTs) and follow-up studies. Primary outcomes were neonatal hypoglycaemia, birthweight and long-term metabolic outcomes. Study quality was assessed using RoB 2.0 and ROBINS-I. Data were synthesized using the IVhet model.

Results: Ten RCTs were included. Metformin was associated with a statistically significant reduction in neonatal hypoglycaemia (OR: 0.65, 95% CI: 0.46–0.92) and lower birthweight (MD: −68.96 g, 95% CI: −108.34 to −29.57). A non-significant trend towards reduced LGA risk was observed. No significant differences in prediabetes, diabetes or insulin resistance were found.

Long-term outcomes in children remain uncertain due to limited and heterogeneous follow-up data.

Conclusions: Metformin appears safe and effective in GDM management, but more data are needed on long-term outcomes.

KEYWORDS

gestational diabetes mellitus (GDM), maternal outcomes, metformin, neonatal outcomes, randomized controlled trials (RCTs), systematic review

Mathilde Louise Saxdorff Brinkmann and Nicoline Josefine Krasilnikoff contributed equally to this work and share first authorship.

1 | INTRODUCTION

Gestational diabetes mellitus (GDM) is increasingly recognized as a significant public health concern, with rising prevalence rates observed globally. The condition is characterized by glucose intolerance diagnosed during pregnancy, leading to a spectrum of complications for both the mothers and their offspring. Obstetric complications associated with GDM include hypertensive disorders, such as preeclampsia and an increased likelihood of caesarean delivery, while neonatal complications may include macrosomia and hypoglycaemia.[1–3] Effective management of GDM is crucial to mitigate these risks and enhance maternal and neonatal outcomes.

Women with a history of GDM face long-term consequences, such as an increased risk of developing type 2 diabetes (T2D) and cardiovascular diseases, while their children are at a heightened risk for obesity and metabolic syndrome later in life.[4–7]

The implications of GDM extend beyond the postpartum period, necessitating ongoing surveillance and preventive strategies for affected families. GDM management strategies in pregnancy consist of self monitoring of blood glucose, and lifestyle modifications including dietary changes focusing on carbohydrates and increased physical activity. However, when these interventions fail to achieve adequate glycaemic control, pharmacological treatments become necessary. Insulin therapy is commonly employed; however, it carries risks such as hypoglycaemia, weight gain and potential adverse effects on maternal mental health, including increased anxiety related to insulin management.[8–10] This underscores the need for alternative treatment options, such as metformin, which has garnered attention for its efficacy in managing blood glucose levels during pregnancy. Metformin enhances insulin sensitivity and reduces hepatic glucose production, and has been associated with weight reduction, making it a viable option for many women with GDM.[11–13]

While metformin has potential benefits, concerns about its safety during pregnancy persist. Metformin crosses the placenta, resulting in comparable maternal and fetal concentrations. While no teratogenic effects have been reported, studies indicate that metformin may influence fetal metabolic programming, with potential long-term consequences for offspring health. In utero exposure to metformin has been linked to a greater prevalence of obesity and increased waist-to-height ratios in children aged 5–10 years. Some studies suggest that while metformin may effectively manage blood glucose levels, the long-term effects on both maternal and offspring health require further investigation.[14–16]

What's new?

- Across ten RCTs, metformin use was associated with a statistically significant reduction in neonatal hypoglycaemia and a modest reduction in birthweight.
- Long-term outcomes for mothers and offspring remain uncertain due to limited and heterogeneous follow-up data.
- Metformin appears to be a safe and effective treatment for GDM, but further research is needed to clarify potential long-term metabolic and developmental effects.

This systematic review and meta-analysis aims to evaluate the safety and efficacy of metformin in managing GDM, focusing on maternal, neonatal and long-term outcomes for both mothers and their offspring. By incorporating recent studies, we aim to provide an updated and comprehensive understanding of the implications of metformin therapy during pregnancy, drawing from randomized controlled trials (RCTs) and their follow-up studies to clarify maternal and neonatal outcomes associated with its use.

2 | METHOD

2.1 | Search strategy

We conducted a systematic review in accordance with the PRISMA Statement[17] to evaluate the effects of metformin on maternal and neonatal outcomes in pregnant women with gestational diabetes. Using the PICO framework (Population, Intervention, Comparison, Outcome), (Appendix Table 1), we have structured our analysis as follows: the population included pregnant women with GDM, the intervention was Metformin treatment, the comparison was standard care or placebo, and the outcomes encompassed both maternal and neonatal health, with primary outcomes being neonatal hypoglycaemia, birthweight and long-term maternal and offspring health. A comprehensive search was performed in PubMed and Embase (Appendix Table 2) to identify relevant RCTs and follow-up studies, with search terms including 'metformin', 'gestational diabetes' and 'maternal outcomes' (Appendix Table 2). We excluded grey literature (e.g. unpublished studies, conference abstracts, theses) from this review, as we specifically focused on published RCTs and their follow-up studies to ensure the highest methodological quality. No filters were applied, and all records were

imported into Covidence for screening. The review is registered in PROSPERO 08.16.2024 (ID=CRD42024576698).

2.2 | Selection criteria

We defined the following eligibility criteria for studies to be included in the review.

2.2.1 | Types of studies

We included RCTs that investigated the use of metformin during pregnancy, with at least 100 participants randomized in the intervention group, focusing on maternal and neonatal outcomes. Follow-up studies assessing long-term outcomes for both mothers and offspring were also included, if they were based on the RCTs included in this review.

2.2.2 | Types of participants

Studies were required to involve pregnant women without pre-existing type 1 or type 2 diabetes, who had not received any antidiabetic medications other than metformin or insulin during pregnancy and who were in their second or third trimester receiving metformin specifically for GDM. We excluded studies involving women with polycystic ovary syndrome (PCOS) or those treated with systemic glucocorticoids, as PCOS represents a separate indication for metformin treatment, often initiated independently of GDM and potentially during the first trimester, which could confound the effects of metformin.

2.2.3 | Types of outcome measures

We included studies reporting on our primary maternal outcomes (e.g. hypertensive disorders, pharmacological treatment for hyperglycaemia, gestational weight gain and mode of delivery) and primary neonatal outcomes (e.g. birthweight, large for gestational age [LGA], small for gestational age [SGA], preterm birth, neonatal hypoglycaemia, neonatal death and stillbirth). Long-term outcomes for both mothers and offspring were considered, including follow-up studies on metabolic health.

However, several of these outcomes were not reported in the included studies. As the aim was to conduct meta-analyses, outcomes needed to be consistently reported across multiple studies. Consequently, only those outcomes with sufficient data were included in the analysis.

2.3 | Data collection and quality assessment

2.3.1 | Selection of studies

Two independent reviewers Mathilde Louise Saxdorff Brinkmann (MLSB) and Nicoline Josefine Krasilnikoff (NJK) screened the titles and abstracts of all unique citations to identify potentially relevant studies for full-text review (Figure 1). Relevant articles were then full text screened based on inclusion and exclusion criteria cf. 2.2 selection criteria. Disagreements regarding study eligibility were resolved through discussion until consensus was achieved. In cases of unresolved disagreement, the opinion of Christina Anne Vinter was sought.

2.3.2 | Data extraction

A standardized data extraction form was created to collect relevant data from all included studies. Data extraction was performed independently by two reviewers MLSB and NJK using Covidence. The reviewers compared their findings and resolved discrepancies through discussion, leading to the creation of a consolidated final form for each study included in the review. Baseline characteristics were verified to ensure comparability across studies (Appendix Table 5). Additionally, data for both our primary and secondary outcomes were extracted.

2.3.3 | Quality assessment

The methodological quality of the included RCTs was assessed using the Risk of Bias 2.0 (RoB 2.0) tool, (Appendix Figure 1), which evaluates the risk of bias across several domains, including randomization, blinding and outcome reporting. For the follow-up studies, we employed the Risk of Bias in Non-randomized Studies-of Interventions (ROBINS-I) tool, (Appendix Figure 1). The assessment allowed us to categorize each study as having low, moderate or high risk of bias. In accordance with recent methodological recommendations, no studies were excluded from synthesis based on their risk-of-bias classification. Instead, we applied the Quality Effects Model, which adjusts study weights

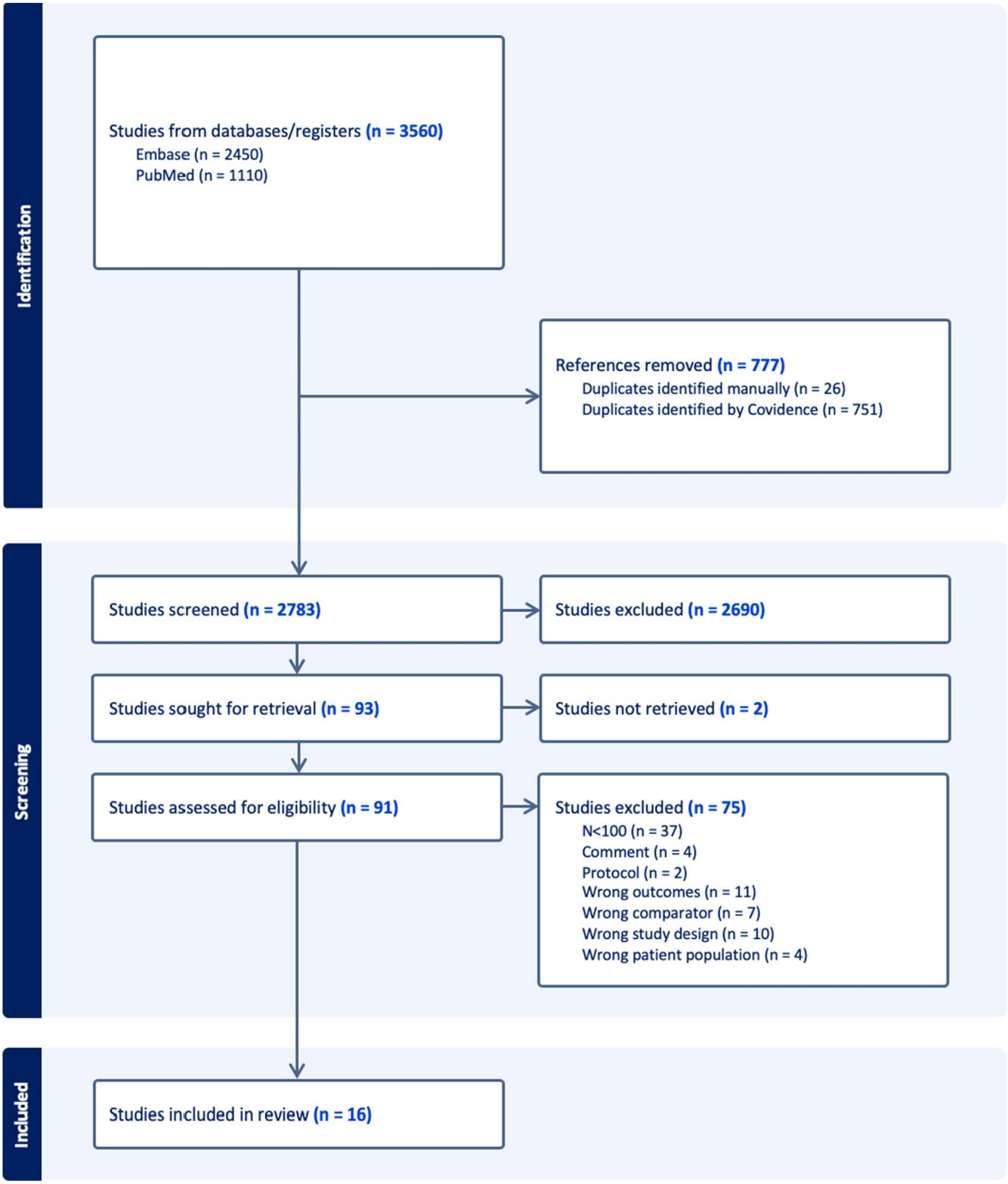

FIGURE 1 Flow diagram of study selection.

based on methodological quality scores. Quality weights (QWT) were calculated by summing the domain-level scores from RoB 2.0 (maximum 5) and ROBINS-I (maximum 7), and RoB 2.0 scores were rescaled to a 0–7 range (multiplied by 7/5) to ensure comparability. These QWT scores were entered into the metan module in Stata (2024 version) for each outcome. Forest plots based on QWT weighting and the primary analyses prior

to sensitivity exclusions are presented in the Appendix Material (Appendix Figures 8–15).

2.4 | Data synthesis

We conducted meta-analyses using the IVhet model in Stata (metan module, 2024 update), which provides more accurate error estimates than traditional random-effects models and avoids underestimating statistical error in the presence of heterogeneity.[18,19] For dichotomous outcomes, we calculated odds ratios (OR) with 95% confidence intervals. Relative risks (RR) were not used, as ORs are now considered more appropriate for meta-analytic synthesis in clinical data.[20] For continuous outcomes, we used mean differences (MD). To improve clinical interpretability, we converted ORs into absolute measures using the logittorisk module in Stata. This yielded baseline risk (p_0), estimated risk with intervention (p_1), risk difference (RD) and number needed to treat or harm (NNT/NNH), based on plausible baseline risks derived from the control group. Heterogeneity was assessed using the I^2 statistic and its associated *p*-value. Small-study effects were evaluated using the DOI plot and LFK index (doiplot module in Stata). Where heterogeneity was suspected due to variation in comparator groups (e.g. placebo vs insulin), we performed sensitivity analyses excluding such studies and reported both pooled and sensitivity results. To account for study quality in the synthesis, we used the Quality Effects model (qeffect option in metan), weighting studies based on their rescaled RoB scores, with RoB 2.0 scores (max 5) rescaled to match ROBINS-I (max 7). All meta-analyses were conducted in Stata 17.

3 | RESULTS

3.1 | Study selection

The systematic review identified a total of 3560 papers through database searches. After removing 26 duplicates identified manually and 751 duplicates identified by Covidence, 2783 unique studies were screened based on their titles and abstracts. Following this initial screening, 2690 studies were excluded based on title and abstract in relation to our PICO (Appendix Table 1), leading to 93 studies that were assessed for full-text retrieval. Of these, 2 could not be retrieved, leaving 91 studies assessed for full-text eligibility. Ultimately, 75 studies were excluded for various reasons: 37 had fewer than 100 participants, 11 reported on incorrect outcomes according to our criteria, 10 had inappropriate study designs, 4 were comments, 7 had wrong comparators, 4 involved incorrect patient populations and 2 were protocols, Consequently, 16 studies were included in the review, comprising 10 RCTs[21–30] and 6 follow-up studies on RCTs[31–36] (Appendix Table 5). A flow diagram illustrating this selection process is presented in Figure 1. Detailed reasons for the exclusion of full-text articles are provided in Appendix Table 3.

3.2 | Study characteristics

The characteristics of the included studies are summarized in Appendix Table 5. In addition, an appendix table provides an overview of which maternal and neonatal outcomes were reported by each study (Appendix Table 4).

3.3 | Risk of Bias and Certainty of Evidence

The risk of bias was assessed using the Risk of Bias 2.0 (RoB 2.0) tool for RCTs and the ROBINS-I (Appendix Figure 1) tool for follow-up studies. No studies were excluded from synthesis based on risk-of-bias ratings.

The certainty of evidence for each outcome was evaluated based on the methodological quality and consistency of the included studies. For outcomes such as neonatal hypoglycaemia and LGA, the evidence was rated as moderate, despite the inclusion of some studies with a high risk of bias. This rating was justified by the consistency of results across most studies, which strengthened confidence in the findings. In contrast, outcomes such as birthweight and maternal insulin sensitivity (HOMA-IR) were associated with lower certainty of evidence. This was due to significant heterogeneity and methodological concerns in certain studies, particularly those identified with a high risk of bias through RoB 2.0 and ROBINS-I. In addition, small-study effects were assessed using DOI plots for the outcomes neonatal hypoglycaemia, LGA, preeclampsia, maternal diabetes and prediabetes. The LFK indices indicated substantial asymmetry in all cases, suggesting potential small-study bias: neonatal hypoglycaemia (−2.06), LGA (−2.41), preeclampsia (−2.53), maternal diabetes (−2.26) and prediabetes (−2.09). These findings are presented in the appendix material (Appendix Figures 20–22).

3.4 | Meta-analysis results

Meta-analyses were conducted for the primary maternal and neonatal outcomes using the IVhet model.[18,19,37–39] Forest plots, pooled effect sizes, heterogeneity measures and absolute risk conversions are presented below. Where relevant, sensitivity analyses were performed to assess the

robustness of the findings and explore the impact of specific studies.

3.4.1 | Large for Gestational Age

Meta-analysis of five studies assessing the risk of delivering a large-for-gestational-age (LGA) infant showed a pooled OR of 0.87 (95% CI: 0.67–1.13), indicating a non-significant trend towards reduced odds with metformin compared to insulin (Appendix Figure 2). Heterogeneity was absent ($I^2 = 0.0\%$, $p = 0.645$). Dunne et al.[22] was excluded from the analysis as the comparator was placebo rather than insulin, making it clinically inconsistent with the other included studies. Prior to this exclusion, the pooled OR was 0.77 (95% CI: 0.55–1.08), with heterogeneity of $I^2 = 37.9\%$, $p = 0.154$.

To enhance clinical interpretability, absolute effect estimates were calculated using the logittorisk module in Stata. The baseline risk for LGA in the insulin group was 17.7%, and the estimated risk with metformin was 7.9%, corresponding to a risk difference (RD) of −9.8 percentage points. This implies that one case of LGA is prevented for approximately every 10th woman treated with metformin rather than insulin (number needed to benefit = 10.2).

3.4.2 | Neonatal hypoglycaemia

Meta-analysis of nine studies assessing the risk of neonatal hypoglycaemia showed a pooled OR of 0.62 (95% CI: 0.40–0.96), indicating a statistically significant reduction in odds with metformin compared to insulin (Figure 2). Heterogeneity was moderate ($I^2 = 63.8\%$, $p = 0.005$). Mir et al.[29] reported unusually high event rates in the insulin group, which contributed disproportionately to the pooled estimate. To assess robustness, a sensitivity analysis was conducted excluding Mir et al.,[29] resulting in a slightly higher OR of 0.65 (95% CI: 0.46–0.92) with reduced heterogeneity ($I^2 = 45.4\%$, $p = 0.077$).

To enhance clinical interpretability, absolute effect estimates were calculated using the logittorisk module in Stata. The baseline risk for neonatal hypoglycaemia in the insulin group was 14.8%, and the estimated risk with metformin was 10.1%, corresponding to a risk difference (RD) of −4.6 percentage points. This implies that one case of neonatal hypoglycaemia is prevented for approximately every 22nd woman treated with metformin rather than insulin (number needed to treat = 22).

Additional estimates across a range of baseline risks are presented in Appendix Figure 16. For example, assuming baseline risks of 10%, 20% or 30%, the corresponding risk

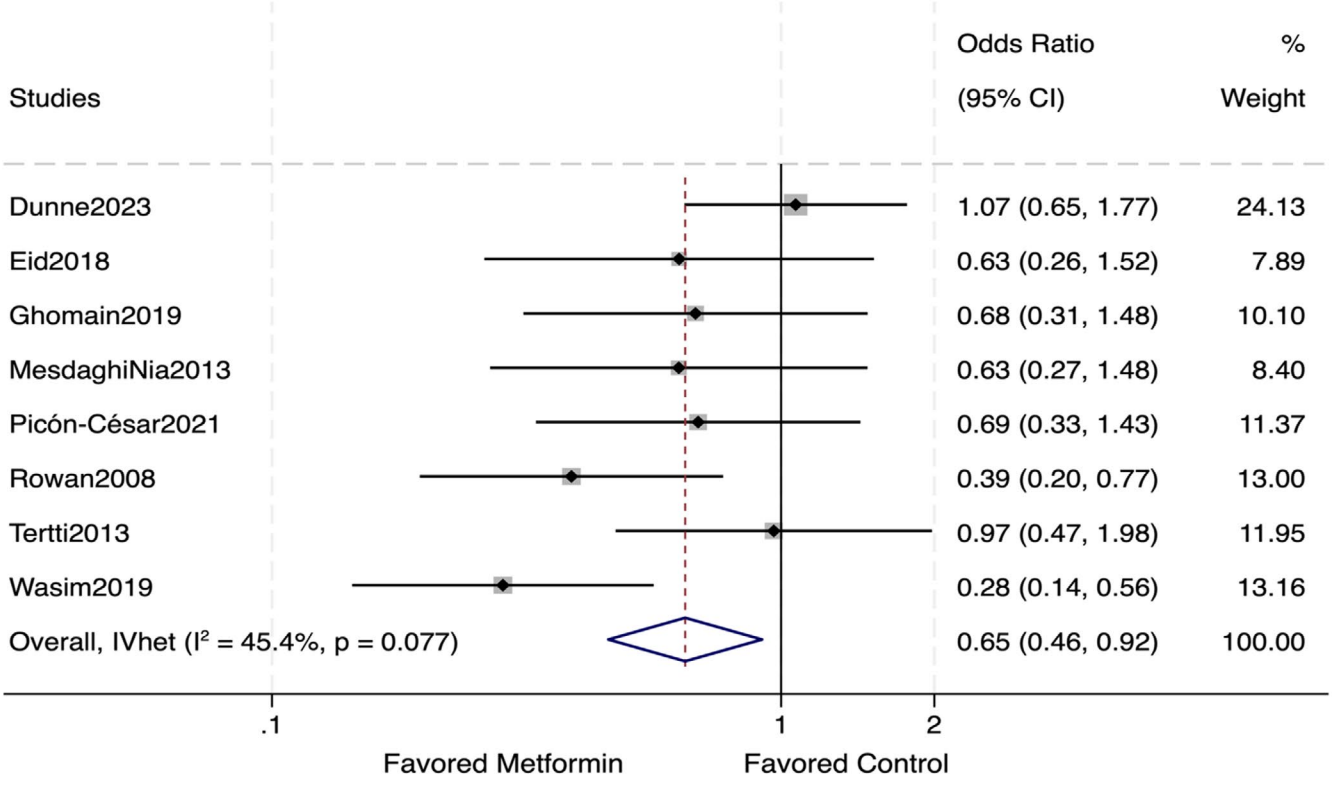

FIGURE 2 Forest plot of odds ratio (OR) for neonatal hypoglycaemia, with 95% confidence intervals and study weights. The diamond represents the pooled estimate.

differences were −3.3, −6.0 and −8.2 percentage points, with NNTs of 31, 17 and 12, respectively.

3.4.3 | Birthweight

Meta-analysis of seven studies (excluding Eid et al.,[24] due to extreme outlier values) showed that metformin was associated with a statistically significant reduction in birthweight compared to insulin, with a pooled mean difference (MD) of −68.96 g (95% CI: −108.34 to −29.57) (Figure 3). Heterogeneity was absent ($I^2 = 0.0\%$, $p = 0.652$). Prior to excluding Eid et al.,[24] the pooled mean difference was −52.3 g (95% CI: −104.6 to 0.0), and heterogeneity was moderate ($I^2 = 48.1\%$, $p = 0.084$). The exclusion of this influential study both increased the precision of the estimate and eliminated heterogeneity, supporting a more consistent effect across studies. These findings suggest that metformin may lead to a modest but consistent reduction in fetal growth compared to insulin in women with GDM.

3.4.4 | Preeclampsia

A meta-analysis of four studies (excluding Dunne et al.,[22] which used placebo instead of insulin) yielded a pooled OR of 0.66 (95% CI: 0.45–0.97), indicating a statistically significant reduction in the risk of preeclampsia with metformin compared to insulin (Appendix Figure 3). Heterogeneity was absent ($I^2 = 0.0\%$, $p = 0.800$). Before excluding Dunne2023, the pooled OR was 0.75 (95% CI: 0.50–1.13), and heterogeneity was slightly higher ($I^2 = 14.8\%$, $p = 0.320$). The exclusion reduced heterogeneity and shifted the effect estimate from non-significant to significant. This suggests that Dunne2023 introduced clinical inconsistency and statistical imprecision due to its different comparator. Absolute effect estimates calculated using logittorisk showed a baseline risk of 9.5% in the insulin group and 8.2% with metformin. This corresponds to a risk difference of −1.3 percentage points, with a number needed to treat (NNT) of 75.5, indicating that one case of preeclampsia may be prevented for every 76th woman treated with metformin rather than insulin.

Appendix Figure 17 presents a full range of baseline risk scenarios. For example, at baseline risks of 10%, 15% and 20%, the risk differences were −3.2, −4.6 and −5.8 percentage points, with corresponding NNBs of 32, 22 and 17, respectively—highlighting increasing absolute benefit at higher baseline risk levels.

3.4.5 | Total fat percent (DEXA) children and total fat (DEXA) children

Meta-analysis of three studies reporting on total fat mass (g) in mid-childhood (7–9 years) showed no statistically significant difference between offspring exposed to metformin versus insulin in utero. The pooled mean difference was −437.6 g (95% CI: −2379.5 to 1504.3),

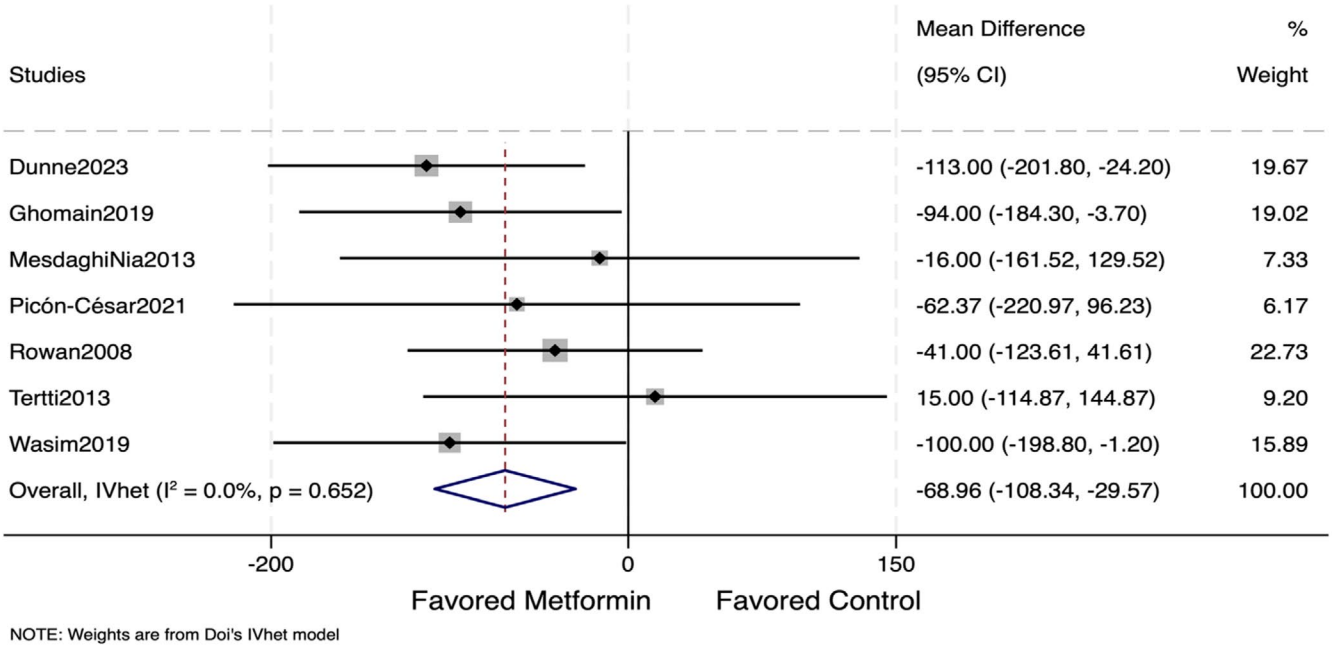

FIGURE 3 Forest plot of mean differences in birthweight, with 95% confidence intervals and study weights. The diamond represents the pooled estimate.

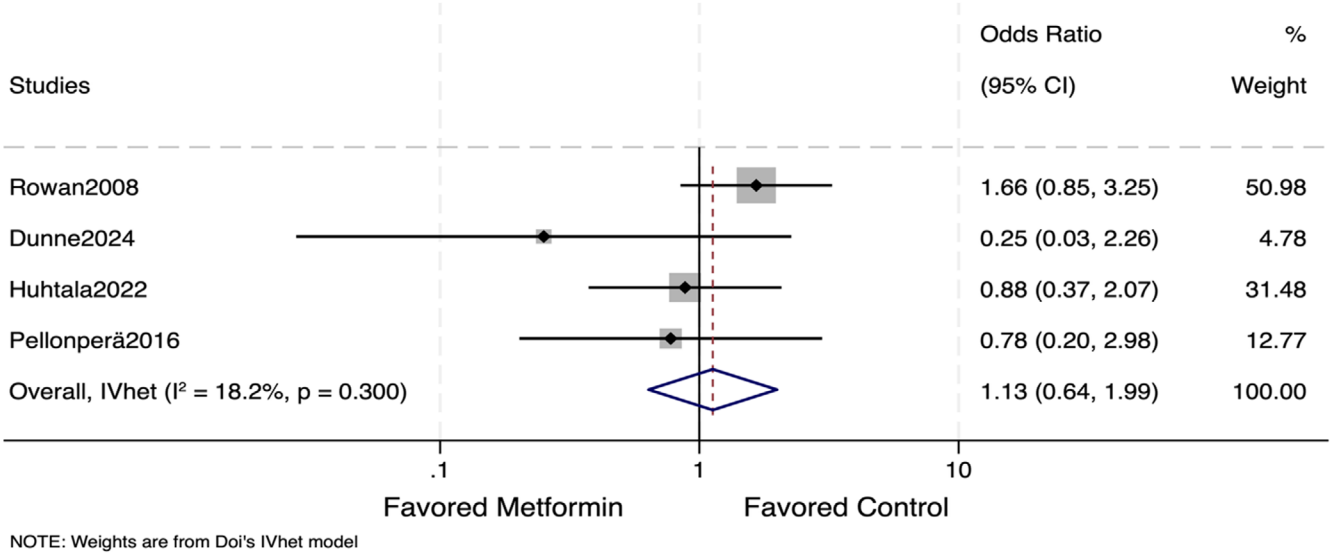

FIGURE 4 Forest plot of OR for maternal diabetes, with 95% confidence intervals and study weights. The diamond represents the pooled estimate.

with substantial heterogeneity ($I^2 = 67.9\%$, $p = 0.044$) (Appendix Figure 4). Weights were calculated using Doi's IVhet model. The high degree of heterogeneity was primarily driven by the Rowan et al.[31] Auckland cohort, which differed markedly in maternal baseline characteristics compared to other populations. The same three studies also reported on total fat percentage. The pooled mean difference was −0.94 percentage points (95% CI: −3.39 to 1.52), again showing no statistically significant difference between groups. Heterogeneity was moderate ($I^2 = 50.5\%$, $p = 0.133$) (Appendix Figure 5). Rowan et al.[36] was excluded from both analyses due to the younger age of offspring at follow-up (2–3 years), which was not comparable to the remaining studies. In addition, the cohort overlapped with Rowan et al.,[31] which reported results separately for the Adelaide and Auckland sites.

3.4.6 | Maternal metabolic long-term outcomes

Meta-analysis of four studies assessing maternal prediabetes showed a pooled OR of 0.91 (95% CI: 0.71–1.17), with no heterogeneity ($I^2 = 0.0\%$, $p = 0.786$) (Appendix Figure 6). For maternal type 2 diabetes, the pooled OR was 1.13 (95% CI: 0.64–1.99) with low heterogeneity ($I^2 = 18.2\%$, $p = 0.300$) (Figure 4). Two studies reported on insulin sensitivity (HOMA-IR), with a pooled mean difference of −0.26 (95% CI: −0.62 to 0.10) and no heterogeneity ($I^2 = 0.0\%$, $p = 0.699$) (Appendix Figure 7). No significant differences were observed in any of the maternal long-term metabolic outcomes.

Appendix Figure 19 presents a full range of baseline risk scenarios for maternal prediabetes, at baseline risks of 10%, 15% and 20%, the risk differences were −0.8, −1.2 and −1.5 percentage points, with corresponding numbers needed to treat (NNTs) of 122, 86 and 68, respectively. For maternal type 2 diabetes, Appendix Figure 18, at baseline risks of 5%, 10% and 15%, the risk differences were +0.6, +1.2 and +1.6 percentage points, with numbers needed to harm (NNHs) of 163, 87 and 62, respectively, indicating small and uncertain absolute effects across the risk spectrum.

4 | DISCUSSION

This systematic review and meta-analyses provide evidence suggesting potential benefits of metformin in managing GDM. Specifically, metformin appears to reduce the risk of neonatal hypoglycaemia and preeclampsia, and is associated with modest reductions in birthweight when compared to insulin. However, these findings should be interpreted cautiously due to the limited number of large, high-quality studies, imprecision in some effect estimates and the methodological necessity to exclude certain studies to reduce heterogeneity. Unlike prior reviews that combined women with GDM and type 2 diabetes (T2D), this review focuses exclusively on GDM, enhancing clinical specificity. While the results indicate promising benefits of metformin, particularly in neonatal outcomes, important gaps remain—especially concerning long-term maternal and offspring health. Further large-scale, well-designed studies with diverse populations and longer follow-up are needed to confirm these findings and establish

the safety profile of metformin as a first-line treatment for GDM.

4.1 | Comparison with previous systematic reviews and meta-analyses

Several previous systematic reviews and meta-analyses have examined the effectiveness of metformin compared to insulin for managing GDM. These studies consistently indicate that metformin is a safe and effective alternative, yielding comparable outcomes in glycaemic control as well as maternal and neonatal health.[3,40,41] Balsells et al.,[42] a systematic review and meta-analysis, provided early evidence supporting metformin's role as a viable alternative to insulin, demonstrating comparable efficacy when used alone or alongside insulin. Farrar et al.,[15] another systematic review and meta-analysis incorporating newer trials previously excluded from earlier reviews, reaffirmed metformin's benefits across a spectrum of pregnancy outcomes, solidifying its acceptance in clinical practice. Butalia et al.[43] conducted a systematic review and meta-analysis including both short- and long-term outcomes of metformin use in pregnant women with GDM and T2D, suggesting potential metabolic benefits for offspring exposed in utero. This aligns with findings from McGrath et al.,[44] a retrospective case–control study reporting no significant differences in infant outcomes between metformin- and insulin-treated mothers. However, Butalia et al. highlighted that the relatively small number of T2D patients and limited long-term data prevent definitive conclusions about similarities or differences in treatment effects between GDM and T2D during pregnancy.

In summary, the aggregated evidence from these systematic reviews and meta-analyses supports the safety and efficacy of metformin in managing GDM, consistent with our findings derived from a more focused population and updated meta-analytic methods.[3,15,40–44]

4.2 | Maternal outcomes

The review has revealed no significant differences in maternal risks such as prediabetes, type 2 diabetes or postpartum insulin resistance between metformin and insulin treatments. All analyses showed no heterogeneity, indicating consistent results across studies. These findings are consistent with previous systematic reviews,[3,15,34] which support metformin's comparable safety profile. Although metformin was associated with a small absolute increase in maternal diabetes risk and a negligible reduction in prediabetes, none of these effects were statistically significant. This suggests that metformin's postpartum metabolic effects are likely modest, and clinically uncertain. The lack of clear benefit on long-term metabolic outcomes indicates that metformin's main advantages may lie in treatment convenience and adherence rather than superior efficacy. While the robustness of the findings is supported by low heterogeneity, the evidence does not establish superiority across all maternal subgroups. Future studies should assess long-term follow-up beyond the early postpartum period to clarify potential subgroup effects and better define metformin's role in maternal metabolic health.

4.3 | Neonatal outcomes

Metformin was associated with a significant reduction in neonatal hypoglycaemia compared to insulin. While the absolute effect may seem modest, neonatal hypoglycaemia is a common and clinically important complication associated with short- and long-term morbidity, including the risk of brain injury. This suggests a potentially meaningful benefit, especially in settings with limited neonatal care resources. This finding aligns with previous meta-analyses,[3,15,35] which have highlighted metformin's potential to reduce neonatal metabolic complications. However, sensitivity to the exclusion of Mir et al.[29] indicates that further validation in diverse populations and clinical contexts is warranted. In contrast, the analysis for large-for-gestational-age infants showed a non-significant trend towards lower risk with metformin. Although absolute estimates suggested a reduction, statistical uncertainty remains, as the confidence interval crossed the threshold for significance. The exclusion of Dunne et al.,[22] which used placebo rather than insulin, improved clinical consistency across studies. Regarding birthweight, metformin was associated with a modest but statistically significant reduction compared to insulin. However, previous findings, including Guo et al.,[3] have noted inconsistencies in this outcome, possibly due to confounding factors such as maternal glycaemic control or population-specific differences. Future research should aim to standardize outcome definitions and reporting and further explore neonatal fat composition and longer term metabolic effects.

4.4 | Offspring body composition and long-term outcomes

The ability to draw meaningful conclusions regarding long-term body composition following prenatal metformin exposure is limited by both the small number of available studies and variability across study populations. Only three studies reported on DEXA-derived outcomes

in mid-childhood (7–9 years), and the pooled results showed no statistically significant differences in total fat mass or fat percentage between metformin- and insulin-exposed offspring. Interpretation is further complicated by substantial heterogeneity observed for total fat mass, partly attributable to differences in baseline characteristics in the Auckland subgroup of Rowan et al.[31]

4.5 | Strengths and limitations of the evidence and the review process

By including only RCTs and their follow-up studies, we aimed to minimize bias and ensure a high level of evidence. A notable strength of our review is the independent screening performed by two blinded reviewers (MLSB and NJK), with third-party involvement to resolve any disagreements or uncertainties. This independent and blinded approach strengthened the reliability of our study selection process. The use of standardized data extraction and quality assessment tools (RoB 2.0 and ROBINS-I) added to the reliability of our findings. Moreover, the involvement of multiple blinded reviewers in data extraction and quality assessment helped reduce potential bias and improve the consistency of our results.

To generate pooled estimates, we used the inverse variance heterogeneity (IVhet) model. This method offers more robust confidence intervals than the traditional random-effects model and reduces the risk of underestimating uncertainty, particularly in the presence of low or moderate heterogeneity. This choice was motivated by current methodological guidance favouring IVhet over conventional approaches.

Effect estimates were expressed as ORs rather than risk ratios (RRs), in line with contemporary recommendations to avoid the statistical limitations associated with RRs in meta-analytic contexts. To enhance clinical interpretability, we used the logittorisk module in Stata to convert ORs into absolute risk differences (RDs) and numbers needed to treat or harm (NNT/NNH) across assumed baseline risks. Small-study effects were explored using the doiplot module, and findings are reported in the Appendix.

However, several limitations must be acknowledged. The presence of high-risk bias in some included studies, particularly regarding missing protocols and incomplete outcome data, raises concerns about the robustness of our conclusions. Some studies exhibited a high risk of bias, which could have introduced reporting biases. For instance, Mir et al.[29] were excluded from the analysis of neonatal hypoglycaemia due to incomplete outcome data, which raised concerns about reporting bias. Similarly, Eid et al.[24] showed potential signs of selective reporting, where certain results may not have been fully disclosed.

We set a threshold of >100 patients in RCTs to improve statistical power and reduce the risk of small-study effects, aiming for more stable and generalizable treatment estimates. However, this decision may have limited our ability to capture potentially meaningful effects observed in smaller trials. Excluding such studies could underestimate treatment benefits or harms, especially for outcomes with low event rates or in specific subgroups. It also increases the risk of selection bias at the review level, as smaller studies with different findings may be systematically left out. While this threshold was chosen to prioritize methodological rigor, its impact on the comprehensiveness of our findings should be acknowledged.

Furthermore, the exclusion of grey literature may have limited our understanding of the broader context of metformin use in GDM management, potentially overlooking relevant insights from unpublished studies. The implications of our findings are significant for clinical practice and public health policy. Given the rising prevalence of GDM and its associated risks, metformin presents a promising alternative to insulin, particularly for women concerned about weight gain and hypoglycaemia.

The analysis was further complicated by variability in study populations, treatment protocols and definitions of outcomes across the included trials. Variations in glycaemic targets, dosage regimens and maternal baseline characteristics likely influenced the observed variability in outcomes such as birthweight and maternal insulin requirements. These issues have been similarly documented in a review by Farrar et al.[15] emphasizing the need for standardization and improved methodological rigor in future studies.

4.6 | Implications of guideline variation for metformin use in GDM

Clinical practice differs due to variation in international guideline recommendations on metformin for GDM. The American College of Obstetricians and Gynecologists (ACOG) and the American Diabetes Association (ADA) recommend insulin as first-line pharmacological treatment, reserving metformin for women who decline, cannot afford or cannot safely use insulin; the ADA also advises caution in women with hypertension, preeclampsia or risk of intrauterine growth restriction. In contrast, the Society for Maternal-Fetal Medicine (SMFM) and the National Institute for Health and Care Excellence (NICE) consider metformin an acceptable first-line option—after 20 weeks' gestation—for women with insufficient glycaemic control on medical nutrition therapy (MNT). The International Federation of Gynecology and Obstetrics (FIGO) also supports its use but outlines scenarios where insulin is preferred. In our meta-analysis, metformin was

associated with lower rates of neonatal hypoglycaemia (OR 0.62, 95% CI 0.40–0.96), modestly reduced birthweight (MD −68.96 g, 95% CI −108.34 to −29.57) and lower risk of preeclampsia (OR 0.66, 95% CI 0.45–0.97) compared with insulin, alongside a non-significant trend towards fewer large-for-gestational-age births. No differences were observed in long-term offspring adiposity or maternal metabolic outcomes, though small-study effects suggested possible publication bias for some outcomes.

Given the variation across guidelines and the limitations of the current evidence base, findings from this systematic review may help guide stakeholders in developing and updating recommendations, but should not dictate treatment decisions. Clinical management should remain individualized, considering patient preferences, clinical context and local resource availability.

4.7 | Future research directions

To address current evidence gaps, future studies should focus on: (i) Standardizing outcome definitions—particularly for birthweight, gestational weight gain and other frequently reported end points. Adoption of the recently developed Core Outcome Set (COS) for reporting of studies on prevention and treatment of GDM, which includes 14 key short-term maternal and neonatal outcomes, can improve consistency in reporting and facilitate data synthesis.[45] (ii) Designing large-scale, multicentre trials with diverse populations to improve generalizability. (iii) Conducting longitudinal studies to assess the long-term health of mothers and their offspring, including metabolic, developmental and cardiovascular outcomes.

5 | CONCLUSION

This systematic review and meta-analysis provides moderate evidence supporting metformin as a viable treatment option for managing GDM. The findings suggest that metformin is associated with favourable maternal and neonatal outcomes, particularly a reduced risk of neonatal hypoglycaemia, without a significant increase in adverse maternal effects. These results are consistent with previous systematic reviews and support metformin as a potential alternative to insulin therapy.

However, caution is warranted. The evidence on long-term outcomes remains limited, and variability in key measures such as birthweight underscores the need for further investigation. Metformin crosses the placenta and exposes the fetus to pharmacologically active concentrations, raising important questions about potential long-term metabolic and developmental effects.

To enhance the certainty and applicability of findings, future studies should prioritize the use of standardized outcome definitions, include larger and more diverse populations, and assess long-term maternal and offspring outcomes. While metformin shows promise, a measured interpretation is essential until more definitive long-term data become available.

FUNDING INFORMATION

This review did not receive any financial or non-financial support from any external funding sources or sponsors.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests related to this review.

DATA AVAILABILITY STATEMENT

All data extracted from the included studies, along with any other materials used in this review, are available upon request. Analytic code and data collection templates are not publicly available.

ARTIFICIAL INTELLIGENCE (AI)

Artificial intelligence (AI) was solely used to assist with wording suggestions and draft formulations, with all content critically reviewed and edited by the authors.

ORCID

Mathilde Louise Saxdorff Brinkmann https://orcid.org/0009-0008-6032-0299
Nicoline Josefine Krasilnikoff https://orcid.org/0009-0009-8425-3892
Mariam Ali Chaaban https://orcid.org/0009-0004-4345-5117
Birgitte Moeller Luef https://orcid.org/0000-0002-8356-5925
Christine Newman https://orcid.org/0000-0003-2387-4109
Fidelma Dunne https://orcid.org/0000-0003-3682-9403
Dorte Moeller Jensen https://orcid.org/0000-0002-3298-9824
Christina Anne Vinter https://orcid.org/0000-0001-5084-6053

REFERENCES

1. Ferrara A. Increasing prevalence of gestational diabetes mellitus: a public health perspective. *Diabetes Care*. 2007;30(Suppl 2):S141-S146.
2. Metzger BE, Lowe LP, Dyer AR, et al. Hyperglycemia and adverse pregnancy outcomes. *N Engl J Med*. 2008;358(19):1991-2002.
3. Guo L, Ma J, Tang J, Hu D, Zhang W, Zhao X. Comparative efficacy and safety of metformin, glyburide, and insulin in treating gestational diabetes mellitus: a meta-analysis. *J Diabetes Res*. 2019;2019:9804708.

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.
5. Tobias DK, Stuart JJ, Li S, et al. Association of History of gestational diabetes with long-term cardiovascular disease risk in a large prospective cohort of US women. *JAMA Intern Med*. 2017;177(12):1735-1742.
6. Kramer CK, Campbell S, Retnakaran R. Gestational diabetes and the risk of cardiovascular disease in women: a systematic review and meta-analysis. *Diabetologia*. 2019;62(6):905-914.
7. Mantzorou M, Papandreou D, Pavlidou E, et al. Maternal gestational diabetes is associated with high risk of childhood overweight and obesity: a cross-sectional study in pre-school children aged 2–5 years. *Medicina (Kaunas)*. 2023;59(3):455.
8. Russell-Jones D, Khan R. Insulin-associated weight gain in diabetes--causes, effects and coping strategies. *Diabetes Obes Metab*. 2007;9(6):799-812.
9. Cheung NW. The management of gestational diabetes. *Vasc Health Risk Manag*. 2009;5(1):153-164.
10. Fraser EE, Ogden KJ, Radford A, et al. Exploring the psychological wellbeing of women with gestational diabetes mellitus (GDM): increased risk of anxiety in women requiring insulin. A prospective longitudinal observational pilot study. *Health Psychol Behav Med*. 2023;11(1):2170378.
11. Reece SW, Parihar HS, LoBello C. Metformin in gestational diabetes mellitus. *Diabetes Spectr*. 2014;27(4):289-295.
12. Shuster DL, Shireman LM, Ma X, et al. Pharmacodynamics of metformin in pregnant women with gestational diabetes mellitus and nonpregnant women with type 2 diabetes mellitus. *J Clin Pharmacol*. 2020;60(4):540-549.
13. Long-term safety, tolerability, and weight loss associated with metformin in the diabetes prevention program outcomes study. *Diabetes Care*. 2012;35(4):731-737.
14. Management of Diabetes in pregnancy: standards of medical care in diabetes-2018. *Diabetes Care*. 2018;41(Suppl 1):S137-s143.
15. Farrar D, Simmonds M, Bryant M, et al. Treatments for gestational diabetes: a systematic review and meta-analysis. *BMJ Open*. 2017;7(6):e015557.
16. Toft JH, Økland I. Metformin use in pregnancy: what about long-term effects in offspring? *Acta Obstet Gynecol Scand*. 2024;103(7):1238-1241.
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
18. Johnson K, Hayen A, Lassere MND, Mengersen K. Bias, coverage, and asymptotic behaviour of random effects meta-analysis: a clinically driven simulation study. *JBI Evid Implement*. 2020;18(4):355-367.
19. Liu Z, Al Amer FM, Xiao M, et al. The normality assumption on between-study random effects was questionable in a considerable number of Cochrane meta-analyses. *BMC Med*. 2023;21(1):112.
20. Doi SA, Furuya-Kanamori L, Xu C, Lin L, Chivese T, Thalib L. Controversy and debate: questionable utility of the relative risk in clinical research: paper 1: a call for change to practice. *J Clin Epidemiol*. 2022;142:271-279.
21. Rowan JA, Hague WM, Gao W, Battin MR, Moore MP. Metformin versus insulin for the treatment of gestational diabetes. *N Engl J Med*. 2008;358(19):2003-2015.
22. Dunne F, Newman C, Alvarez-Iglesias A, et al. Early metformin in gestational diabetes: a randomized clinical trial. *JAMA*. 2023;330(16):1547-1556.
23. Tertti K, Ekblad U, Koskinen P, Vahlberg T, Rönnemaa T. Metformin vs. insulin in gestational diabetes. A randomized study characterizing metformin patients needing additional insulin. *Diabetes Obes Metab*. 2013;15(3):246-251.
24. Eid SR, Moustafa RSI, Salah MM, et al. Is metformin a viable alternative to insulin in the treatment of gestational diabetes mellitus (GDM)? Comparison of maternal and neonatal outcomes. *Egyptian Pediatric Association Gazette*. 2018;66(1):15-21.
25. Mesdaghinia E, Samimi M, Homaei Z, Saberi F, Moosavi SG, Yaribakht M. Comparison of newborn outcomes in women with gestational diabetes mellitus treated with metformin or insulin: a randomised blinded trial. *Int J Prev Med*. 2013;4(3):327-333.
26. Ghomian N, Vahed SHM, Firouz S, Yaghoubi MA, Mohebbi M, Sahebkar A. The efficacy of metformin compared with insulin in regulating blood glucose levels during gestational diabetes mellitus: a randomized clinical trial. *J Cell Physiol*. 2019;234(4):4695-4701.
27. Picón-César MJ, Molina-Vega M, Suárez-Arana M, et al. Metformin for gestational diabetes study: metformin vs insulin in gestational diabetes: glycemic control and obstetrical and perinatal outcomes: randomized prospective trial. *Am J Obstet Gynecol*. 2021;225(5):517.e1-571.e17.
28. Dunne F, Newman C, Alvarez-Iglesias A, et al. Postnatal cardiometabolic health after metformin use in gestational diabetes: a secondary analysis of the emerge trial. *J Clin Endocrinol Metab*. 2025;110(5):e1566-e1572.
29. Mir S, Saleem R, Saghir F, Zaib Mir S, Iqbal A, Manzoor A. To compare metformin vs insulin in gestational diabetes in terms of neonatal hypoglycemia. *Pakistan Journal of Medical and Health Sciences*. 2021;15(11):2928-2929.
30. Wasim T, Shaukat S, Javaid L, Mukhtar S, Amer W. Comparison of metformin and insulin for management of gestational diabetes mellitus: a randomized control trial. *Pakistan Journal of Medical & Health Sciences*. 2019;13:823-827.
31. Rowan JA, Rush EC, Plank LD, et al. Metformin in gestational diabetes: the offspring follow-up (MiG TOFU): body composition and metabolic outcomes at 7-9 years of age. *BMJ Open Diabetes Res Care*. 2018;6(1):e000456.
32. Pellonperä O, Rönnemaa T, Ekblad U, Vahlberg T, Tertti K. The effects of metformin treatment of gestational diabetes on maternal weight and glucose tolerance postpartum--a prospective follow-up study. *Acta Obstet Gynecol Scand*. 2016;95(1):79-87.
33. Huhtala M, Nikkinen H, Paavilainen E, et al. Comparison of glucose metabolism and anthropometry in women with previous gestational diabetes treated with metformin vs. insulin: 9-year follow-up of two randomized trials. *Acta Obstet Gynecol Scand*. 2022;101(5):514-523.
34. Paavilainen E, Niinikoski H, Parkkola R, et al. Metformin versus insulin for gestational diabetes: adiposity variables and adipocytokines in offspring at age of 9 years. *Diabetes Res Clin Pract*. 2023;202:110780.
35. Paavilainen E, Tertti K, Nikkinen H, et al. Metformin versus insulin therapy for gestational diabetes: effects on offspring anthropometrics and metabolism at the age of 9 years: a follow-up study of two open-label, randomized controlled trials. *Diabetes Obes Metab*. 2022;24(3):402-410.

36. Rowan JA, Rush EC, Obolonkin V, Battin M, Wouldes T, Hague WM. Metformin in gestational diabetes: the offspring follow-up (MiG TOFU): body composition at 2 years of age. *Diabetes Care*. 2011;34(10):2279-2284.
37. Doi SAR, Furuya-Kanamori L. Selecting the best meta-analytic estimator for evidence-based practice: a simulation study. *Int J Evid Based Healthc*. 2020;18(1):86-94.
38. Doi SAR, Furuya-Kanamori L, Thalib L, Barendregt JJ. Meta-analysis in evidence-based healthcare: a paradigm shift away from random effects is overdue. *Int J Evid Based Healthc*. 2017;15(4):152-160.
39. Senn S. Trying to be precise about vagueness. *Stat Med*. 2007;26(7):1417-1430.
40. Shah MK, Shah MA, Shah SD, Shah PK, Patel K, Gupta M. Comparison of neonatal and maternal outcomes of anti-diabetic drugs in the treatment of gestational diabetes mellitus: findings from Bayesian network meta-analysis. *J Family Med Prim Care*. 2022;11(5):1610-1618.
41. Bidhendi Yarandi R, Amiri M, Ramezani Tehrani F, Behboudi-Gandevani S. Effectiveness of antidiabetic agents for treatment of gestational diabetes: a methodological quality assessment of meta-analyses and network meta-analysis. *J Diabetes Investig*. 2021;12(12):2247-2258.
42. Balsells M, García-Patterson A, Solà I, Roqué M, Gich I, Corcoy R. Glibenclamide, metformin, and insulin for the treatment of gestational diabetes: a systematic review and meta-analysis. *BMJ*. 2015;350:h102.
43. Butalia S, Gutierrez L, Lodha A, Aitken E, Zakariasen A, Donovan L. Short- and long-term outcomes of metformin compared with insulin alone in pregnancy: a systematic review and meta-analysis. *Diabet Med*. 2017;34(1):27-36.
44. McGrath RT, Glastras SJ, Scott ES, Hocking SL, Fulcher GR. Outcomes for women with gestational diabetes treated with metformin: a retrospective, case-control study. *J Clin Med*. 2018;7(3):50.
45. Egan AM, Bogdanet D, Griffin TP, et al. A core outcome set for studies of gestational diabetes mellitus prevention and treatment. *Diabetologia*. 2020;63(6):1120-1127.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Brinkmann MLS, Krasilnikoff NJ, Chaaban MA, et al. Metformin safety during pregnancy in women with gestational diabetes mellitus: A systematic review and meta-analysis of maternal, neonatal and long-term outcomes. *Diabet Med*. 2025;00:e70173. doi:10.1111/dme.70173