diabetesresearchandclinicalpractice231(2026)113070

Contents lists available at ScienceDirect

Diabetes Research and Clinical Practice

journal homepage: www.journals.elsevier.com/diabetes-research-and-clinical-practice

Prediction of time to insulin initiation in gestational diabetes mellitus: a secondary analysis of the EMERGE trial

Yueyun Zhu [a,*], Alberto Alvarez-Iglesias [b,g], Aoife M. Egan [c,e], Andrew Smyth [b,c,d], Christine Newman [b,c,d,g,j], David Aguilar Macias [b], Declan Devane [f,h], Fidelma Dunne [b,c,d,g,j], Martin O'Donnell [b,c,d], Paula O'Shea [c], Paddy Gillespie [i,j], Andrew J. Simpkin [a]

[a] School of Mathematical and Statistical Sciences, University of Galway, Galway, Ireland
[b] HRB Clinical Research Facility Galway, University of Galway, Galway, Ireland
[c] College of Medicine Nursing and Health Sciences, University of Galway, Galway, Ireland
[d] Galway University Hospital, Newcastle Road, Galway, Ireland
[e] Mayo Clinic, Rochester, MN, USA
[f] School of Nursing and Midwifery, University of Galway, Galway, Ireland
[g] Institute for Clinical Trials, University of Galway, Galway, Ireland
[h] Health Research Board – Trials Methodology Research Network, University of Galway, Galway, Ireland
[i] School of Business and Economics, University of Galway, Galway, Ireland
[j] Clinical Trials Network (CTN) Diabetes, Galway, Ireland

ARTICLE INFO

Keywords:
Gestational diabetes mellitus
Insulin therapy
Random survival forest
Decision curve analysis

ABSTRACT

Aims: Gestational diabetes mellitus (GDM) is a global health problem. Insulin therapy is recommended when lifestyle management fails to control blood glucose. We aim to predict the time to insulin initiation for women with GDM.

Methods: A random survival forest (RSF) model was developed to predict the time to insulin initiation among 413 women from the EMERGE trial, analysed separately for placebo and metformin groups. Maternal characteristics and early glucose data (collected during the two weeks after randomisation) were used as predictors. Decision curve analysis was performed to assess the net benefit of the model compared with default strategies.

Results: The RSF model had a concordance index (C-index) of 0.71 (95 % CI: 0.64–0.77), time-dependent AUC $\geq$0.70 and Brier score $\leq$0.2 for the placebo group, and a C-index of 0.72 (95 % CI: 0.64–0.80), time-dependent AUC $\geq$0.75 and Brier score $\leq$0.2 for the metformin group. The decision curve analysis showed the RSF provided a higher net benefit for both groups compared to default strategies across clinically relevant threshold probabilities.

Conclusions: The RSF model effectively identified women at high risk of requiring insulin. The decision curve analysis could help clinicians to balance insulin initiation decisions. However, prospective validation is needed to confirm the generalisability of the model.

1. Introduction

Gestational diabetes mellitus (GDM) is defined as carbohydrate intolerance resulting in hyperglycaemia with onset during pregnancy [1]. It is one of the most common complications of pregnancy and potentially increases the risk of several adverse maternal outcomes such as excessive weight gain, caesarean birth and preeclampsia, as well as adverse neonatal outcomes including macrosomia, birth injuries, hypoglycaemia and respiratory distress [2–4].

The initial management of GDM typically involves lifestyle intervention with a focus on nutritional and exercise changes [5]. For women who do not achieve desired glucose levels after initial lifestyle intervention, pharmacologic therapy is required. Insulin has traditionally been the first-choice drug for GDM [5,6]. The timely initiation of insulin therapy is crucial in reducing the risk of adverse maternal and neonatal outcomes [7]. Glucose monitoring plays a pivotal role in the timely initiation of insulin therapy for women with GDM [8]. In this study of women with GDM, we use glucose data collected in the first two weeks

* Corresponding author at: School of Mathematical and Statistical Sciences, University of Galway, Galway H91 TK33, Ireland.
E-mail address: yueyun.zhu@universityofgalway.ie (Y. Zhu).

https://doi.org/10.1016/j.diabres.2025.113070
Received 4 November 2025; Received in revised form 17 December 2025; Accepted 20 December 2025
Available online 21 December 2025

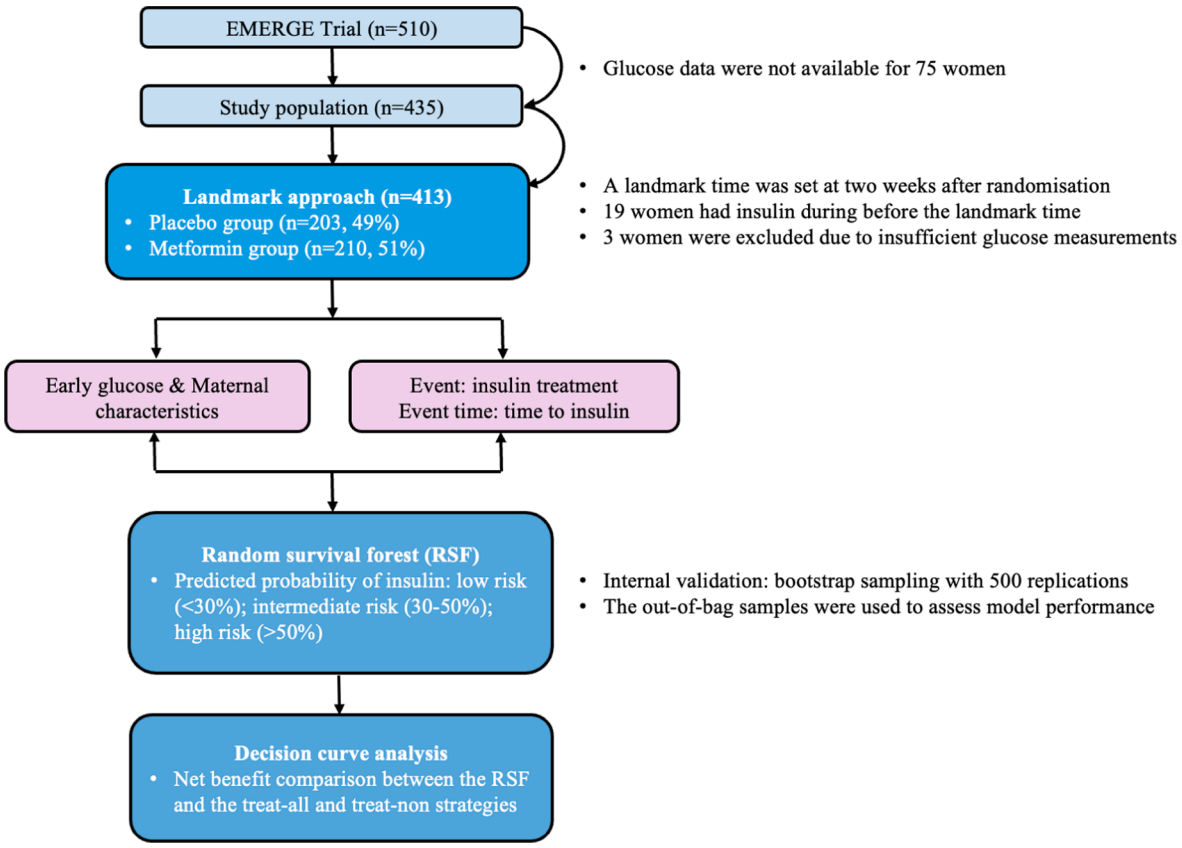

Fig. 1. Study workflow A total of 435 women were included in this cohort study. Among them, 413 women who were at risk at the landmark time (i.e, two weeks after randomisation) were included in the landmark analysis. A RSF model was employed to predict the time to insulin initiation for placebo and metformin groups, respectively, using predictors from early glucose (i.e. first two weeks of recording) and maternal characteristics. Internal validation was performed using bootstrap sampling with 500 iterations. The predicted risk groups were defined based on the predicted probability of insulin initiation: low (<30 %), intermediate (30–50 %) and high (>50 %). A decision curve analysis was used to compare the net benefit of RSF with treat-all and treat-none strategies for guiding insulin treatment decisions.

after randomisation to predict insulin initiation in later pregnancy.

Previous studies [9–11] have identified risk factors associated with insulin requirement in women with GDM. These studies found that body mass index (BMI), family history of diabetes, pre-existing hypertension, a prior history of GDM and the number of abnormal results on the 100-g/3-h oral glucose tolerance test (OGTT) are positively correlated with the need for insulin. Several studies [12–14] have employed logistic regression models to predict the probability of insulin therapy at the individual level. However, it is also clinically important to estimate the time to insulin initiation [15]. In this study, we use the database from the Early Metformin in Gestational Diabetes (EMERGE) trial [16,17]. Prior work [16] has used the Kaplan-Meier method to compare the time to insulin initiation between women in the placebo and metformin groups. However, the Kaplan-Meier analysis does not account for multiple predictors simultaneously, nor does it provide individualised risk estimates.

Our study aims to develop prediction models which include multiple potential risk factors to predict the time to insulin initiation for women in the placebo and metformin groups separately. In particular, we aim to use early glucose recordings and maternal characteristics to predict insulin initiation in later pregnancy. The model and its prediction result can help guide clinical decision-making and support timely intervention.

2. Subjects, materials and methods

2.1. Data source

Data for this study were obtained from the EMERGE trial. EMERGE was a phase 3, parallel-group, two-site, superiority, randomised, double-blind, placebo-controlled trial of metformin (in addition to usual care) in women diagnosed with GDM [16]. The trial design and primary outcome of EMERGE have been reported elsewhere [16,17]. In summary, from June 2017 to September 2022, 535 singleton pregnancies (in 510 women) with GDM (defined by the World Health Organization 2013 criteria (1)) were randomised to receive either metformin or a matched placebo from randomisation to delivery. Randomisation occurred on the day of screening or up to 7 days post-screening. Participants received a 75-g/2-h OGTT at the screening time. Laboratory tests, such as glycated haemoglobin (HbA1c) and cholesterol, were performed at screening, 32 and 38 gestational weeks and before delivery. Insulin was started if two or more home glucose readings between clinic visits were outside the prespecified glucose targets (fast level, ≤5.1 mmol/L; 1-hour postprandial level, ≤7mmol/L) and could not be explained by transient factors such as sleep, stress or infection. Moreover, participants were instructed by the diabetes team to perform seven-point glucose monitoring daily, measuring glucose levels before and one hour after each meal, as well as at bedtime. Glucose data were unavailable for some participants, with records successfully downloaded from the glucometer for 435 of 510 participants.

2.2. Study design and population

A landmark [18] approach was used in the study design. The landmark approach involves setting a specific time point (i.e., landmark time) and using data available up to this time to predict future outcomes. In this study, the landmark time was set at two weeks after

randomisation, which allowed sufficient time to observe initial lifestyle intervention effects and ensured an adequate risk set for modelling. Prior to the landmark time, 19 women out of 435 (4 %) who had already required insulin therapy and 3 women who had insufficient glucose measurements (fewer than four records) were excluded. Consequently, 413 women (out of 435) who had sufficient glucose records and had not required insulin initiation from randomisation up to the landmark time constituted the at-risk cohort, see Fig. 1. For the at-risk cohort, data collected up to the landmark time were used to predict the time to insulin initiation later in pregnancy. Note that a longer landmark period would substantially reduce the risk set. For instance, using a three-week period excluded 59 women (14 %) from the analysis.

2.3. Predictors

Maternal and pregnancy data were collected from the medical records and the EMERGE trial database [16,17], including randomisation group (placebo or metformin), maternal age at randomisation (years), gestational week at randomisation (weeks), body mass index (BMI, kg/m^2) at randomisation, previous GDM, smoking status (never, former, current), first pregnancy status, OGTT (fasting, 1 h and 2 h postprandial) and HbA1c at randomisation. Additionally, early glucose data were selected from randomisation up to the landmark time (i.e., two weeks after randomisation). This two-week window allowed sufficient glucose measurements to characterise early glycaemic response. The individual mean, maximum and standard deviation of early glucose data were calculated for each woman. Moreover, the glucose data were collected repeatedly over time and can be treated as functional data. Functional principal component analysis (FPCA) [19] was applied to early glucose trajectories, providing a set of functional principal component scores (FPC scores) for each participant. For clarity, we refer to the first three FPC-scores as *score1*, *score2* and *score3*, respectively. The FPC scores serve as a compact and powerful summary of the key features of the glucose trajectory for each participant. For example, *score1* reflected the patient-specific glucose level during the two-week landmark window, with higher values indicating higher glucose levels. A more detailed explanation of FPC-scores is provided in the Supplementary Material.

2.4. Prediction model

A random survival forest (RSF) [20] was employed to predict the time to insulin initiation for women with GDM. Compared to the commonly used Cox proportional hazards models, RSF can capture complex, non-linear relationships between variables and survival outcomes. A detailed explanation of RSF models is provided in the Supplementary Material. A total of 16 predictors were used in the RSF model, including gestational week at randomisation, maternal age at randomisation, BMI at randomisation, previous GDM, smoking status, first time pregnancy, three OGTT values at randomisation (fasting, 1 h and 2 h postprandial), HbA1c at randomisation, the mean, maximum and standard deviation of early glucose data, and three FPC-scores (*score1, score2, score3*). For comparison, Cox models using the same predictors were also developed. The RSF and Cox models were applied to women in the placebo and metformin groups separately.

2.5. Evaluation

To prevent overfitting, internal validation was performed using bootstrap sampling with 500 replications. The out-of-bag samples (approximately 36.8 % of the data) were used to assess model performance based on the following metrics. The concordance index (C-index) reflects the model's ability to rank individuals based on their risk of experiencing an event, ranging from 0.5 (no better than random prediction) to 1 (perfect prediction). The time-dependent Brier score measures the average squared difference between the predicted survival probability and actual outcome over a specified prediction horizon, with lower values indicating better predictive performance. The time-dependent area under the receiver operating characteristic curve (AUC) measures the model's ability to distinguish between patients at higher and lower risk over a specified prediction horizon, with higher values indicating better predictive accuracy.

Table 1
Baseline characteristics of patients at risk.

Characteristics	Overall N = 413	Placebo N = 203	Metformin N = 210	Missing	p-value
Randomisation week (weeks)	25.2 ± 4.1	25.2 ± 4.1	25.2 ± 4.0	0	0.71
Age (years)	34.7 ± 4.5	34.5 ± 4.4	34.8 ± 4.6	0	0.63
Gestational age at delivery (weeks)	39.1 ± 1.5	39.2 ± 1.5	39.1 ± 1.5	0	0.41
Race				0	0.05
White	339 (82 %)	162 (80 %)	177 (84 %)		
Irish Traveller	7 (2 %)	1 (0 %)	6 (3 %)		
Other	67 (5 %)	40 (6 %)	27 (3 %)		
Smoking status				0	0.14
Current	23 (6 %)	14 (7 %)	9 (4 %)		
Former	175 (42 %)	77 (38 %)	98 (47 %)		
Never	215 (52 %)	112 (55 %)	103 (49 %)		
BMI (kg/m^2)	30.2 ± 6.0	30.5 ± 5.4	30.0 ± 6.5	24	0.10
First pregnancy	100 (24 %)	47 (23 %)	53 (25 %)	0	0.62
Prior macrosomia	77 (19 %)	36 (18 %)	41 (20 %)	0	0.64
Prior GDM	97 (23 %)	48 (24 %)	49 (23 %)	0	0.94
Prior caesarean	115 (28 %)	58 (29 %)	57 (27 %)	0	0.75
Prior preeclampsia	42 (10 %)	21 (10 %)	21 (10 %)	0	0.91
HbA1c at randomisation (mmol/mol)	32.8 ± 3.4	32.9 ± 3.5	32.8 ± 3.3	6	0.96
OGTT (mmol/L)					
Fasting glucose	5.2 ± 0.5	5.2 ± 0.5	5.2 ± 0.4	0	0.70
1-hour postprandial glucose	9.5 ± 1.9	9.7 ± 1.9	9.3 ± 1.9	0	0.02
2-hour postprandial glucose	7.2 ± 1.6	7.2 ± 1.5	7.1 ± 1.6	0	0.43
Insulin required	174 (42 %)	99 (49 %)	75 (36 %)	0	0.01
Insulin age (weeks)	28.7 ± 4.3	28.7 ± 4.1	28.7 ± 4.7	0	0.67
Birth weight (grams)	3451 ± 524	3524 ± 525	3381 ± 514	0	0.01
1-minute APGAR score				5	0.67
0 ~ 3	2 (0 %)	1 (0 %)	1 (0 %)		
4–6	9 (2 %)	6 (3 %)	3 (1 %)		
7 ~ 10	397 (97 %)	194 (97 %)	203 (98 %)		
5-minute APGAR score 7 ~ 10	411 (100 %)	202 (100 %)	209 (100 %)	2	–

Continuous variables are represented as mean ± SD. Categorical variables are represented as n (%). Wilcoxon rank sum test and Pearson's Chi-squared test are used for continuous and categorical variables, respectively. Randomisation week is the gestational week at randomisation (i.e. shortly after GDM diagnosis); Gestational age is the age (weeks) at delivery; BMI, body mass index; HbA1c, haemoglobin A_{1c}; OGTT, oral glucose tolerance test; Insulin age (weeks) is the gestational week at insulin initiation for women who require insulin treatment.

2.6. Statistical analysis

Descriptive statistics were presented as proportions for categorical

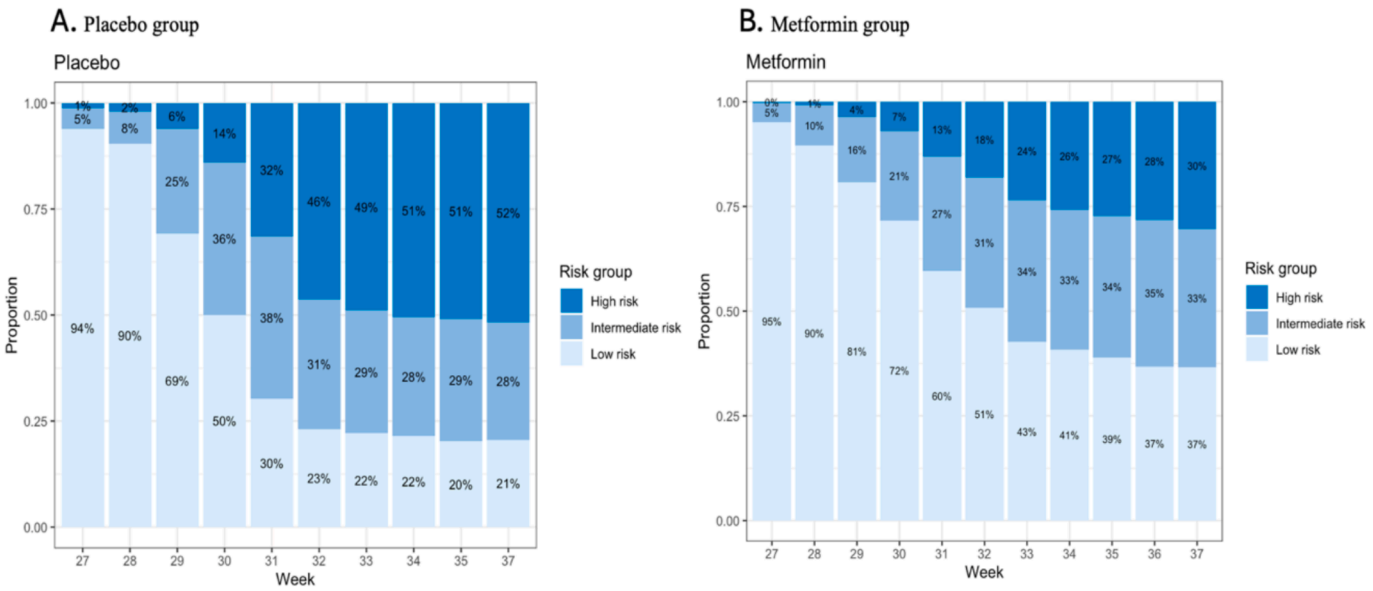

Fig. 2. The proportion of women in different risk groups across later gestational weeks The predicted risk groups were defined based on the predicted probability of insulin initiation from the random survival forest model: low (<30 %), intermediate (30–50 %) and high (>50 %). The x-axis is the gestational weeks after 27, and y-axis represents the proportion of different risk groups.

variables and as mean ± standard deviation for continuous variables. Comparisons between placebo and metformin groups were conducted using the chi-square test for categorical variables and Wilcoxon rank-sum test for continuous variables.

The RSF model was employed to predict the probability of insulin initiation after the landmark time until delivery. Based on the predicted probabilities in each replication of bootstrap sampling, women were stratified into three risk groups: low (<30 %), intermediate (30–50 %) and high (>50 %). Moreover, a decision curve analysis [21] was used to quantify the clinical net benefit of the prediction model compared with default strategies that treat all patients or none. Let n be the sample size and p_t denote the threshold probability, i.e. the minimum risk to advise insulin initiation. The net benefit is defined by

$$Netbenefit = \frac{Truepositive}{n} - \frac{Falsepositive}{n} * \frac{p_t}{1-p_t}$$

We performed the decision curve analysis at the end of pregnancy (around week 39). Women who had a predicted risk higher than p_t and eventually required insulin therapy were defined as true positives, whereas those with a predicted risk higher than p_t but who did not need insulin therapy were defined as false positives. Based on different p_t, the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were also calculated. All analyses were performed by R statistical software (v4.4.2).

3. Results

3.1. Patient characteristics

Table 1 presents the baseline characteristics of women in the placebo group (N = 203, 49 %) and the metformin group (N = 210, 51 %). The fasting and 2-hour postprandial OGTT glucose levels were similar between two groups, whereas 1-hour postprandial OGTT glucose levels were significantly different, with 9.7 ± 1.9 for the placebo and 9.3 ± 1.9 for the metformin (p = 0.02). Importantly, more women in the placebo group (n = 99, 49 %) received insulin therapy compared to the metformin group (n = 75, 36 %; p < 0.01). This finding is consistent with the known efficacy of metformin in reducing insulin requirements. For women who required insulin initiation, the mean gestational week of initiation was 28.7 ± 4.1 in the placebo group and 28.7 ± 4.7 in the metformin group, with no significant difference between groups (p = 0.67).

3.2. Predictive performance in the placebo group

The proportion of women in different risk groups indicates nearly half were predicted to be at high risk of requiring insulin after week 31, see Fig. 2A. In terms of model performance, the C-index of the RSF for the placebo group was 0.71 (95 % CI: 0.64–0.77), which was higher than the Cox model (0.69, 95 % CI: 0.62–0.76). Fig. 3B illustrates that the RSF model provided better predictive performance than the Cox model, with a higher time-dependent AUC (above 0.70 throughout most of the pregnancy) and a lower time-dependent Brier score (below 0.2 throughout most of the pregnancy). The AUC began to decline after week 27, coinciding with the period when more patients started insulin therapy (Fig. 3A). However, the overall AUC remained above 0.70, indicating that the model had good discriminative performance during pregnancy.

3.3. Predictive performance in the metformin group

The proportion of women in different risk groups indicates nearly one third were predicted to be at high risk after week 34, see Fig. 2B. In terms of model performance, the C-index of the RSF for the metformin group was 0.72 (95 % CI: 0.64–0.80), while the C-index of the Cox model was 0.70 (95 % CI: 0.64–0.79). Fig. 3C presents the time-dependent AUC and time-dependent Brier score, which illustrates that the RSF outperformed the Cox with higher AUC values and lower Brier scores. The AUC of RSF for the metformin group was generally above 0.75 throughout most of the pregnancy, whereas the Brier score of RSF was lower than 0.2 over the pregnancy. Consequently, the RSF for the metformin group demonstrated strong discrimination throughout pregnancy.

3.4. Decision curve analysis

Fig. 3 presents the decision curves for the placebo and metformin groups at the end of pregnancy (around week 39). For the placebo group, the RSF model achieved a net benefit of 0.17 at a threshold

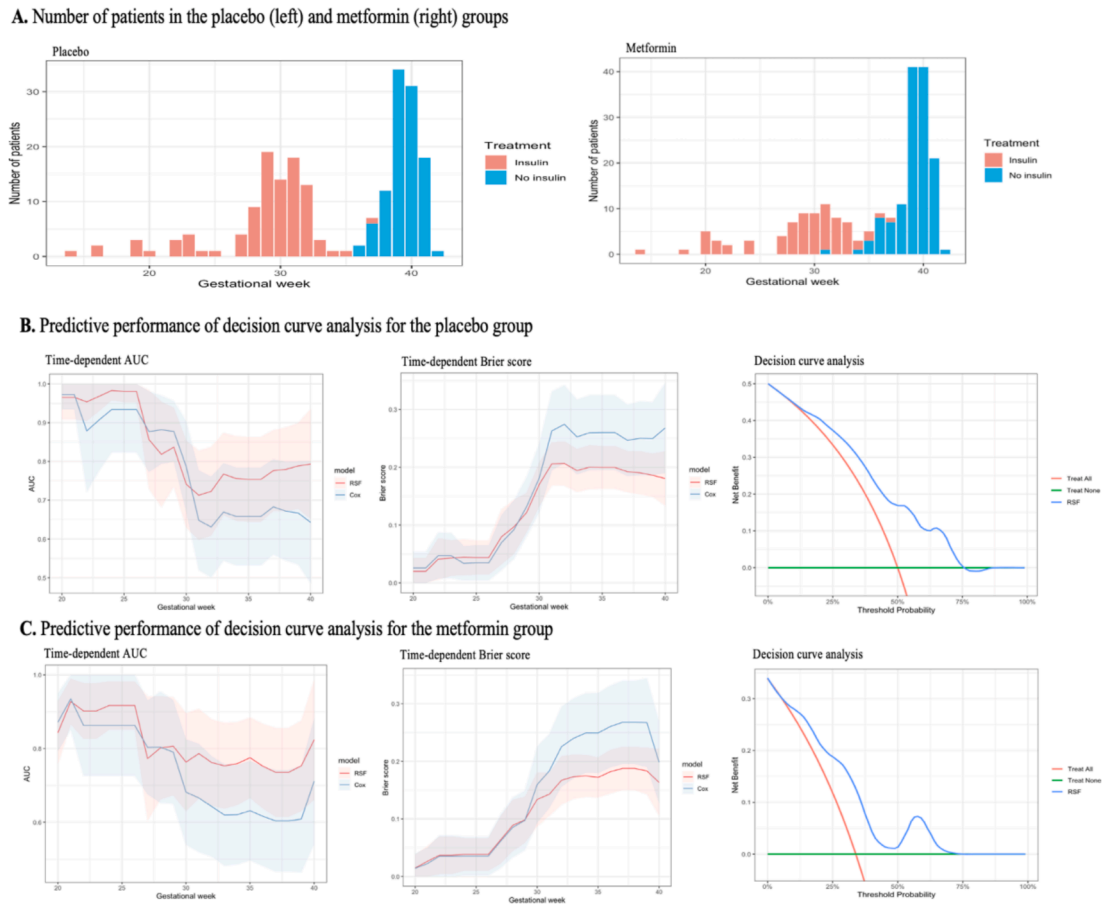

Fig. 3. Distribution of women with and without insulin treatment, predictive performance and decision curve analysis Panel A: Number of patients in the placebo (left) and metformin groups (right) who required insulin (red) and those who did not receive insulin (blue) at each gestational week. Panel B: Time-dependent AUC (left), time-dependent Brier score (middle) and a decision curve at the end of pregnancy (right) for the placebo group. Panel C: Time-dependent AUC (left), time-dependent Brier score (middle) and a decision curve at the end of pregnancy (right) for the metformin group. A higher time-dependent AUC and a lower Brier score indicate better predictive performance. For the decision curves, the x-axis represents the threshold probability, and the y-axis represents the net benefit. The blue line shows the RSF model; the red and green lines represent the reference "treat-all" and "treat-none" strategies, respectively. Values higher on the y-axis indicate that a strategy yields more true-positive insulin initiations without increasing false positives. A net benefit of 0.10, for example, translates to 10 additional correctly identified women per 100 without unnecessary early insulin initiation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

probability of 50 %, whereas the treat-none and treat-all strategies had net benefits of 0. This indicates that, compared to treating all women with insulin and treating none with insulin, the RSF model yielded a benefit equivalent to 17 additional correctly identified women per 100 at that threshold, accounting for the weighting of false positives. For the metformin group, the RSF model yielded a net benefit of 0.13 at a threshold probability of 27 %, whereas the treat-none and treat-all strategies had net benefits of approximately 0. This suggests that the RSF model provided an equivalent of 13 additional correct identifications per 100 at that threshold, with false positives weighted by the threshold probability, compared to treating all women and not treating any women with insulin therapy. Overall, the RSF model yielded a higher net benefit than the default strategies across clinically relevant threshold probabilities. The decision curve plots in Fig. 3 indicates that clinical decision-making thresholds should be tailored separately for the placebo and metformin groups.

3.5. *Threshold probability selection*

Table 2 presents the classification results using a range of threshold probabilities for the placebo and metformin groups. For the placebo group, Table 2 suggests that a cutoff between 0.4 and 0.5 would be suitable, with sensitivity decreasing from 0.85 to 0.72 and specificity increasing from 0.50 to 0.66. For the metformin group, a cutoff between 0.3 and 0.4 would be preferable, with sensitivity decreasing from 0.84 to 0.67 and specificity increasing from 0.48 to 0.63. As a result, a higher decision threshold should be used for the placebo group to achieve a balanced trade-off between sensitivity and specificity, compared to the metformin group. This was because more women in the placebo group required insulin therapy.

4. Discussion

In this secondary analysis of the EMERGE trial, a RSF model was applied to predict the time to insulin initiation for the placebo and metformin groups, separately. Overall, the RSF demonstrated a high level of predictive performance, with a C-index of 0.71 (95 % CI: 0.64–0.77) for the placebo group and 0.72 (95 % CI: 0.64–0.80) for the metformin group. The time-dependent AUC from the RSF was above 0.70, and the time-dependent Brier score was below 0.20 throughout the pregnancy for both the placebo and metformin groups. Moreover, the decision curve analysis indicated that the RSF provided an overall higher net benefit than the treat-all and treat-none strategies across clinically relevant threshold probabilities.

Previous studies have used logistic regression models to identify risk factors associated with the need for insulin [9–13]. The main strength of

Table 2
Threshold probabilities for insulin initiation prediction for the placebo and metformin groups, respectively.

Placebo				
Threshold	Sens	Spec	PPV	NPV
0.1	1 (0.96–1)	0.02 (0–0.1)	0.48 (0.39–0.58)	0.81 (0–1)
0.2	0.96 (0.87–1)	0.15 (0.02–0.31)	0.51 (0.41–0.63)	0.83 (0.56–1)
0.3	0.9 (0.79–1)	0.34 (0.15–0.51)	0.56 (0.43–0.68)	0.8 (0.62–1)
0.4	0.83 (0.67–0.95)	0.51 (0.33–0.71)	0.61 (0.49–0.76)	0.77 (0.62–0.91)
0.5	0.72 (0.51–0.87)	0.65 (0.47–0.81)	0.66 (0.52–0.8)	0.72 (0.56–0.86)
0.6	0.54 (0.3–0.76)	0.77 (0.59–0.92)	0.69 (0.55–0.84)	0.64 (0.5–0.8)
0.7	0.31 (0.13–0.55)	0.88 (0.72–0.97)	0.72 (0.52–0.91)	0.58 (0.47–0.71)
Metformin				
Threshold	Sens	Spec	PPV	NPV
0.1	0.99 (0.93–1)	0.18 (0.06–0.33)	0.4 (0.29–0.5)	0.98 (0.87–1.01)
0.2	0.95 (0.81–1)	0.35 (0.21–0.5)	0.44 (0.34–0.56)	0.93 (0.77–1)
0.3	0.84 (0.61–1)	0.49 (0.33–0.65)	0.47 (0.36–0.6)	0.85 (0.69–1)
0.4	0.64 (0.4–0.87)	0.65 (0.48–0.83)	0.5 (0.35–0.67)	0.77 (0.65–0.91)
0.5	0.44 (0.24–0.68)	0.8 (0.63–0.94)	0.56 (0.37–0.79)	0.73 (0.63–0.85)
0.6	0.3 (0.11–0.5)	0.91 (0.8–0.98)	0.66 (0.43–0.88)	0.71 (0.61–0.81)
0.7	0.17 (0–0.37)	0.96 (0.89–1)	0.69 (0.45–0.88)	0.68 (0.58–0.78)

At the end of the pregnancy, if the predicted probability of requiring insulin is larger than the specified threshold, the patient was classified as "insulin required"; otherwise, the patient was classified as "no insulin therapy". Values in parentheses indicate 95% confidence intervals.
Abbreviations: Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value.

our study is that we used early glucose data (i.e., two weeks after randomisation) and maternal characteristics to predict the time to insulin initiation in later pregnancy based on the RSF model. The prediction results allow clinicians to assess the risk of insulin requirement at an early stage and predict the gestational week at which insulin is likely to be initiated. These results enable clinicians to make informed early insulin decisions, ultimately improving maternal and neonatal outcomes.

Indeed, based on the predicted probabilities of insulin initiation from the RSF model, women were stratified into low (<30 %), intermediate (30–50 %) and high (>50 %) risk groups. In the placebo group, nearly half of the women were predicted to be at high risk after week 31, whereas in the metformin group, approximately one third were predicted to be at high risk after week 34. This closely aligned with the actual clinical treatment decisions, where 49 % of women in the placebo group and 36 % in the metformin group received insulin treatment eventually. To better guide clinical management, women at low risk are advised to continue standard care through diet and lifestyle interventions; women at intermediate risk should have their review brough forward (e.g. by one week), reinforce nutrition and glucose monitoring, begin insulin education and consent early to allow prompt initiation if control drifts; women at high risk warrant strong consideration of insulin initiation, particularly when glucose readings are trending above target levels [22,23]. However, it is important to note that our model is intended to support the decision-making process, while the final treatment decision should still be based on current glucose recordings and clinical assessment.

Establishing safe and effective criterion for pharmacotherapy initiation for GDM is important, as undertreatment may increase GDM-related complications [24,25]; whereas overtreatment may result in greater financial burden for patients [26] and potential adverse effects of medications [27]. Our decision curve analysis compared the RSF model with the treat-none (representing undertreatment) and treat-all (representing overtreatment) strategies. It is observed that the RSF model had an overall higher net benefit than the default strategies across clinically relevant threshold probabilities. This suggests that the RSF offers an opportunity to achieve an optimal balance between undertreatment and overtreatment.

However, it is challenging to determine a single cutoff for insulin initiation, because binary classification models are inherently sensitive to small changes in the threshold. The American College of Obstetricians and Gynecologists also states that there is no conclusive evidence for a specific threshold value at which medical therapy should be started [28]. In our study, we found that for the placebo group, a cutoff between 0.4 and 0.5 provided a reasonable balance between sensitivity (decreasing from 0.85 to 0.72) and specificity (increasing from 0.50 to 0.66), whereas a cutoff between 0.3 and 0.4 would be preferable for the metformin group, with sensitivity from 0.84 to 0.67 and specificity from 0.48 to 0.63. Furthermore, the decision threshold should be tailored separately for the placebo and metformin groups, since more women in the placebo group required insulin therapy compared to the metformin group.

Our study has some limitations. The first is the lack of external validation. However, repeated glucose measures among women with GDM are not readily available, and future studies using continuous glucose monitoring (CGM) or timed finger stick testing data are required to validate our work. Moreover, with the advent of more accurate and low-cost continuous glucose monitors, such data should be easier to collect, and those studies will benefit from the results found here. It should be also noted that the treatment thresholds used in EMERGE are not necessarily applicable across different populations, potentially limiting external validation. Second, our study does not investigate the clinical outcomes of patients who required insulin therapy. Further analysis, such as evaluating maternal and neonatal outcomes for women in different risk groups, could be performed. Third, the model was developed using data from a single randomised trial with specific inclusion criteria, which may limit generalisability to broader clinical populations. Fourth, the landmark approach requires patients to complete two weeks of glucose monitoring, which may not be feasible in all clinical settings.

In conclusion, we employed a RSF model to predict the time to insulin initiation for women with GDM in the placebo and metformin groups separately. The model included predictors from early glucose data (i.e., two weeks after randomisation) and maternal characteristics, demonstrated a high level of predictive performance and effectively identified the cohort at high risk of requiring insulin. Moreover, the decision curve analysis was performed to compare the RSF model with the treat-all and treat-none strategies, and the RSF model showed higher net benefit across clinically relevant threshold probabilities. Our study could help clinicians identify women at high risk of requiring insulin treatment early in pregnancy, facilitate timely intervention and guide clinical decision-making. However, external validation and clinical implementation studies are needed before routine clinical use.

CRediT authorship contribution statement

Yueyun Zhu: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Alberto Alvarez-Iglesias:** Writing – review & editing. **Aoife M. Egan:** Writing – review & editing. **Andrew Smyth:** Writing – review & editing. **Christine Newman:** Writing – review & editing. **David Aguilar Macias:** Writing – review & editing. **Declan Devane:** Writing – review & editing, Funding acquisition. **Fidelma Dunne:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization. **Martin O'Donnell:** Writing – review & editing, Funding acquisition. **Paula O'Shea:** Writing

– review & editing. **Paddy Gillespie:** Writing – review & editing. **Andrew J. Simpkin:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Funding

This study was supported by HRB of Ireland, coordinated by the HRB-Clinical Research Facility Galway and Research Ireland grant 19/FFP/7002.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.diabres.2025.113070.

References

[1] World Health O. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy. In: Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy; 2013. p. 62.
[2] Reece EA, Leguizamón G, Wiznitzer A. Gestational diabetes: the need for a common ground. Lancet 2009;373:1789–97.
[3] O'Sullivan EP, Avalos G, O'reilly M, et al. Atlantic Diabetes in Pregnancy (DIP): the prevalence and outcomes of gestational diabetes mellitus using new diagnostic criteria. Diabetologia 2011;54:1670–5.
[4] Metzger BE, Coustan DR, Trimble ER. Hyperglycemia and adverse pregnancy outcomes. Clin Chem 2019;65:937–8.
[5] American Diabetes Association Professional Practice C, American Diabetes Association Professional Practice C. 15. Management of diabetes in pregnancy: Standards of Medical Care in Diabetes—2022. Diab Care 2022;45:S232–43.
[6] American DA. 9. Pharmacologic approaches to glycemic treatment: standards of medical care in diabetes—2020. Diabetes Care 2020;43:S98–110.
[7] Harrison RK, Cruz M, Wong A, Davitt C, Palatnik A. The timing of initiation of pharmacotherapy for women with gestational diabetes mellitus. BMC Pregn Childbirth 2020;20:773.
[8] Li A, Brackenridge A. The role of continuous glucose monitoring in pregnancy. Obst Med 2022;15:6–10.
[9] Bakiner O, Bozkirli E, Ozsahin K, Sariturk C, Ertorer E. Risk factors that can predict antenatal insulin need in gestational diabetes. J Clin Med Res 2013;5:381.
[10] Sapienza AD, Francisco RPV, Trindade TC, Zugaib M. Factors predicting the need for insulin therapy in patients with gestational diabetes mellitus. Diabetes Res Clin Pract 2010;88:81–6.
[11] Zhang Y, Shao J, Li F, Xu X. Factors in gestational diabetes mellitus predicting the needs for insulin therapy. Int J Endocrinol 2016;2016:4858976.
[12] Eleftheriades M, Chatzakis C, Papachatzopoulou E, et al. Prediction of insulin treatment in women with gestational diabetes mellitus. Nutr Diabetes 2021;11:30.
[13] Zaccara TA, Mikami FCF, Paganoti CF, Francisco RPV, Costa RA. Predicting insulin use among women with gestational diabetes diagnosed in oral glucose tolerance test. BMC Pregnancy Childbirth 2023;23:410.
[14] Barnes RA, Wong T, Ross GP, et al. A novel validated model for the prediction of insulin therapy initiation and adverse perinatal outcomes in women with gestational diabetes mellitus. Diabetologia 2016;59:2331–8.
[15] Du R, Li L. Estimating the risk of insulin requirement in women complicated by gestational diabetes mellitus: a clinical nomogram. Diab Metab Synd Obesity 2021:2473–82.
[16] Dunne F, Newman C, Alvarez-Iglesias A, et al. Early metformin in gestational diabetes: a randomized clinical trial. J Am Med Assoc 2023;330:1547–56.
[17] Dunne F, Newman C, Devane D, et al. A randomised placebo-controlled trial of the effectiveness of early metformin in addition to usual care in the reduction of gestational diabetes mellitus effects (EMERGE): study protocol. Trials 2022;23:795.
[18] Paige E, Barrett J, Stevens D, et al. Landmark models for optimizing the use of repeated measurements of risk factors in electronic health records to predict future disease risk. Am J Epidemiol 2018;187:1530–8.
[19] Ramsay JO, Silverman BW. Principal components analysis for functional data. Funct Data Anal 2005:147–72.
[20] Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. 2008.
[21] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.
[22] Oostdam N, van Poppel MNM, Wouters MGAJ, van Mechelen W. Interventions for preventing gestational diabetes mellitus: a systematic review and meta-analysis. J Womens Health 2011;20:1551–63.
[23] Brown J, Alwan NA, West J, et al. Lifestyle interventions for the treatment of women with gestational diabetes. Cochrane Database Syst Rev 2017.
[24] Hartling L, Dryden DM, Guthrie A, Muise M, Vandermeer B, Donovan L. Benefits and harms of treating gestational diabetes mellitus: a systematic review and meta-analysis for the US Preventive Services Task Force and the National Institutes of Health Office of Medical applications of Research. Ann Intern Med 2013;159:123–9.
[25] Poolsup N, Suksomboon N, Amin M. Effect of treatment of gestational diabetes mellitus: a systematic review and meta-analysis. PLoS One 2014;9:e92485.
[26] Kolu P, Raitanen J, Rissanen P, Luoto R. Health care costs associated with gestational diabetes mellitus among high-risk women–results from a randomised trial. BMC Pregnancy Childbirth 2012;12:71.
[27] Yogev Y, Ben-Haroush A, Chen R, Rosenn B, Hod M, Langer O. Undiagnosed asymptomatic hypoglycemia: diet, insulin, and glyburide for gestational diabetic pregnancy. Obstet Gynecol 2004;104:88–93.
[28] ACOG Practice Bulletin No. 190 Summary: Gestational Diabetes Mellitus. Obstet Gynecol 2018;131:406-408.