

## Research Article

# Assessing the Relationship between Gestational Glycemic Control and Risk of Preterm Birth in Women with Type 1 Diabetes: A Joint Modeling Approach

Resmi Gupta <sup>1</sup>, Jane C. Khoury,<sup>1</sup> Mekibib Altaye,<sup>1</sup> Roman Jandarov,<sup>2</sup> and Rhonda D. Szczesniak <sup>1</sup>

<sup>1</sup>Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital, Cincinnati, Ohio, USA

<sup>2</sup>Department of Biostatistics, University of Cincinnati, Cincinnati, Ohio, USA

Correspondence should be addressed to Resmi Gupta; [resmi.gupta@cchmc.org](mailto:resmi.gupta@cchmc.org)

Received 24 February 2020; Revised 19 May 2020; Accepted 10 June 2020; Published 24 June 2020

Academic Editor: Ulrike Rothe

Copyright © 2020 Resmi Gupta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Background.** Characterizing maternal glucose sampling over the course of the entire pregnancy is an important step toward improvement in prediction of adverse birth outcome, such as preterm birth, for women with type 1 diabetes mellitus (T1DM). **Objectives.** To characterize the relationship between the gestational glycemic profile and risk of preterm birth using a joint modeling approach. **Methods.** A joint model was developed to simultaneously characterize the relationship between a longitudinal outcome (daily blood glucose sampling) and an event process (preterm birth). A linear mixed effects model using natural cubic splines was fitted to predict the longitudinal submodel. Covariates included mother's age at last menstrual period, age at diabetes onset, body mass index, hypertension, retinopathy, and nephropathy. Various association structures (value, value plus slope, and area under the curve) were examined before selecting the final joint model. We compared the joint modeling approach to the time-dependent Cox model (TDCM). **Results.** A total of 16,480 glucose readings over gestation (range: 50-260 days) with 32 women (28%) having preterm birth was included in the study. Mother's age at last menstrual period and age at diabetes onset were statistically significant ( $\beta = 1.29$ , 95% CI 1.10, 1.72;  $\beta = 0.84$ , 95% CI 0.62, 0.98) for the longitudinal submodel, reflecting that older women tended to have higher mean blood glucose and those with later diabetes onset tended to have a lower mean blood glucose level. The presence of nephropathy was statistically significant in the event submodel ( $\beta = 2.29$ , 95% CI 1.05, 4.48). Cumulative association parameterization provided the best joint model fit. The joint model provided better fit compared to the time-dependent Cox model (DIC (JM) = 19,895; DIC (TDCM) = 19,932). **Conclusion.** The joint model approach was able to simultaneously characterize the glycemic profile and assess the risk of preterm birth and provided additional insights and a better model fit compared to the time-dependent Cox model.

## 1. Background

Self-monitoring of blood glucose plays a significant role in reduction of perinatal mortality and morbidity in pregnant women with type 1 diabetes mellitus (T1DM) [1, 2]. Tracking blood glucose over the entire pregnancy has provided insight into the effects of poor maternal glucose control on various neonatal outcomes, including preterm birth [3]. Recent studies have shown that the rate of both spontaneous and indicated preterm birth is increased in women with T1DM [4].

The rate of preterm birth ranged from 22% to 45% in women with T1DM, which is greater by four- to eightfold over the frequency of preterm delivery in pregnancies which are not complicated by diabetes [5]. Preterm babies have been shown to have a higher risk of morbidity, mortality, and poor neurological outcome later in life [6]. In addition to the overall increased rate, poor glycemic control was associated with both spontaneous and indicated preterm birth, for mothers with T1DM [7]. Hence, characterizing the glucose profile together with estimating the rate of maternal glucose

sampling over the course of the entire pregnancy is an important step along the path to predict the risk of preterm birth outcome and inform T1DM care management practices.

Regulation of glucose is a dynamic process that varies in response to meals, insulin dosage as well as gestational period [8]. In recent years, advancement in technology has improved diabetes self-management practices allowing monitoring of blood glucose on both a programmed and continuous basis. The maternal glucose profile may be considered measurements of an almost continuous-time monitoring process whose outputs are samples of functions or curves. Each curve accounts for blood glucose oscillations that occur throughout the day and perhaps at the same time on different days throughout pregnancy. The longitudinal measurements of the blood glucose level are subject to measurement error as most other self-reported and machine-recorded data, but measurements may be viewed as a noisy version of the true glucose profile.

A common aim in follow-up studies is to characterize the relationship between longitudinal measurements and the event outcomes to gain a better understanding of the risk of an event such as preterm delivery. The time-dependent Cox model [9] is prevalent in the literature for modeling this kind of association. However, the time-dependent Cox model assumes that the covariates are external and are independent of the event mechanism [10, 11]. Moreover, another strong assumption in the time-dependent Cox model is that the time-dependent covariates are measured without error.

In recent years, a joint model has been shown to be an efficient approach to simultaneously model longitudinal data and an observed outcome, often yielding more accurate, informative prediction than traditional models [12, 13]. The joint model approach constructs two submodels, longitudinal and event time, and links the two models by subject-specific random effect terms. The early development of joint models for longitudinal and survival data was primarily motivated from HIV/AIDS clinical trials to model longitudinal CD4 counts and survival data. Classical models such as the linear mixed effects model for longitudinal data and the Cox model for time-to-event data do not consider dependencies between these two different data types (longitudinal and time-to-event data). By simultaneously modeling these two data types into a single model, joint models can infer the dependence and association between the longitudinal and time-to-event outcomes, in order to better assess the effect of a treatment, to quantify uncertainty, and to provide accurate predictions of outcomes. Excellent expositions of joint models have been provided by Wulfsohn and Tsiatis [14], Tsiatis and Davidian [15], De Gruttola and Tu [16], Wang and Taylor [17], Henderson et al. [18], and Brown and Ibrahim [19]. More recent work on joint models includes Rizopoulos et al. [20], Wu et al. [21], Ye et al. [22], Huang et al. [23], and Wu et al. [24].

In addition to being able to characterize the association between longitudinal and time-to-event outcomes using the joint model, dynamic prediction has also gained increasing momentum in clinical research. Because dynamic predictions are individualized and have the ability to appropriately account for possible nonlinearity in each individual longitu-

dinal trajectory, better prospective treatment decisions may become available. In contrast to the linear mixed model, the joint model could provide individualized risk prediction based on the availability of longitudinal information up to that point with the assumption that the individual is event-free until that time.

In the current study, we utilize the aforementioned advantages of the joint model in the shared-parameter framework to characterize the relationship between maternal glucose profile in the longitudinal submodel and preterm birth outcome in the event submodel. Our overall objective is to improve the prediction using the joint model approach while characterizing the relationship between maternal glycemic profiles and preterm birth among women with T1DM. We hypothesize that the joint model would provide better model fit compared to the conventional approach, the time-dependent Cox model. The final joint model will be chosen based on the model fit statistics for different association structures between the maternal glycemic profiles and the risk of preterm birth. Additionally, following Rizopoulos-proposed Monte Carlo approach [25], we will estimate the risk of preterm birth and illustrate how it can be dynamically updated, given that the subject-specific glycemic profiles were available up to the time of prediction.

## 2. Methods

*2.1. Study Design and Participant Characteristics.* The study methods and cohort characteristics have been described in detail elsewhere [26]. Briefly, women with confirmed diagnosis of T1DM who were either pregnant or planning a pregnancy were recruited and enrolled in our Diabetes in Pregnancy study. The women were prospectively followed over gestation as part of a 17-year interdisciplinary program of diabetes in pregnancy between 1978 and 1995 conducted at the University of Cincinnati Medical Center. The study subjects were managed with intensive insulin therapy, involving a split mixed-dose regimen of three to four injections per day using short- and intermediate-acting insulin combined with dietary regulation. After 1981, women were instructed to check blood glucose concentrations 6–8 times a day: fasting, preprandial (before each meal), 90 min postprandial (after each meal), at bedtime, and occasionally at 3:00 AM. The current analysis included women who used a reflectance meter through pregnancy and delivered a singleton live fetus beyond 20 weeks of gestation. The glucose measurements recorded between gestational days 50 and 260 were included in the study. Data from profiles corresponding to neonatal death within 28 days of delivery or presence of a major congenital malformation were excluded from this analysis. Birthweight was measured within the first hour of delivery using an electronic scale (Toledo Scale, Worthington, Ohio). Preterm delivery was defined as delivery prior to 37-week gestation as a result of either spontaneous preterm labor, or preterm premature rupture of membranes, or any other cause classified by a study perinatologist as indicated preterm delivery. Glycosylated hemoglobin A1c concentration was measured every 4 weeks throughout pregnancy and prior to delivery using Isolab column chromatography.

Mother's age at last menstrual period, age at diabetes onset, body mass index, white classification (widely used to assess maternal and fetal risk complicated by diabetes), chronic hypertension, retinopathy, and nephropathy were included as predictors in the initial longitudinal (glucose trajectory) and event (preterm birth) models. Bivariate analysis indicated that mother's age, age at diabetes onset, and the presence of nephropathy were statistically significant in both longitudinal and event models and will be included in the joint model and time-dependent Cox modeling phase. Glucose recordings were available for all patients on the day the event (birth) occurred.

**2.2. Statistical Methods.** Each joint model consists of two linked submodels, a mixed effects model to fit longitudinal blood glucose monitoring and a survival model to fit preterm birth data. The longitudinal outcome, glucose recordings (log-transformed scale) from the  $i^{\text{th}}$  woman and  $j^{\text{th}}$  pregnancy observed at gestation time  $t_{ij}$  ( $i = 1 \dots n, j = 1 \dots n_{ij}$ ) can be expressed as

$$\text{Glucose recording}_{ij} = X_{ij}^T(\theta) + f(t_{ij}) + Z_{ij}^T b_i + \varepsilon_{ij}, \quad (1)$$

where parametric effects are represented for traditional covariates  $X_{ij}$  and  $p \times 1$  parameter vector  $\theta$ ;  $f(\cdot)$  is a nonparametric smooth function evaluated at time  $t_{ij}$ ;  $Z_{ij}$ , and  $b_i$  corresponding to the design matrix and subject-specific  $q_i \times 1$  parameter vectors for random effects;  $\varepsilon_{ij}$  represents the measurement error corresponding to the observation at time  $t_{ij}$ . The mean response function  $f(t_{ij})$  in equation (1) below can be estimated using the natural cubic spline which can be expressed as:  $f(t_{ij}) = \sum_{r=0}^{22} B_r \lambda_r(t_{ij})$  where  $B_r$  are the parameter coefficients of the expansion with basis functions  $\lambda_r(t_{ij})$ . In order to capture the nonlinear subject specific fluctuations in glucose recordings, we included the natural cubic spline that expands the time effects into a B-spline basis matrix. Based on the Bayesian information criteria [27], we selected a total of 22 knots, which was necessary to estimate the individual dynamic prediction for risk of preterm birth (discussed in the next section). Including the fixed-effect covariates, maternal age at last menstrual period, age at diabetes onset, and spline functions, the above equation can be written as

$$\begin{aligned} \text{Glucose recording}_{ij} &= m_i(t) + \varepsilon_i(t) = (\beta_0 + b_{i0}) \\ &+ (\beta_1 + b_{i1}) B_r \left( (\text{gestational day})_{ij} + \lambda_1 \right) \\ &+ (\beta_2 + b_{i2}) B_r \left( (\text{gestational day})_{ij} + \lambda_2 \right) \\ &+ \dots \dots + (\beta_{22} + b_{i22}) B_r \left( (\text{gestational day})_{ij} + \lambda_{22} \right) \\ &+ \beta_{23}(\text{age})_i + \beta_{24}(\text{age at diabetes onset})_i + \varepsilon_{ij}. \end{aligned} \quad (2)$$

In equation (1),  $B_r$  represent the fixed effects part of the spline coefficient,  $r = 0, 1, \dots, 22$ . We assumed that the ran-

dom spline coefficients are mutually independent and distributed as  $b_{ik} \sim N_{23}(0, \sigma_{ik})$ ,  $k = 0, 1, \dots, 22$ , and measurement error  $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ .

We linked the event model, probability of preterm birth to the glucose measurements through the random effects. Let  $h_i$  denote the event endpoint, which is the binary indicator of preterm birth. The probability of preterm birth is linked to the longitudinal process of glucose measurements as

$$h_i[t] = lt^{l-1} \exp \{ \Omega_i^T(\phi) + \alpha m_i(t) \}, \quad (3)$$

where  $\Omega_i^T$  denotes the vector of fixed effects covariates, such as the presence of nephropathy at baseline;  $m_i(t)$  is the true values of the underlying glucose monitoring processes, with  $\alpha$  being the parameter characterizing the association between longitudinal glucose recording process and the PB outcome. Because of model complexity we specified the Weibull distributional form for the baseline hazard function. Equation (3) monotonically increases with time if  $l > 1$ , decreases if  $l < 1$ , and the exponential hazard remains constant if  $l = 1$ .

In equation (3), the risk of preterm birth depends on the current value of the blood glucose level. However, since for each patient the blood glucose follows a trajectory in time, it is also reasonable to consider a joint model in which the risk depends on both the current true value of the blood glucose trajectory and the slope of the true blood glucose trajectory at that time. The event model then becomes

$$h_i[t] = lt^{l-1} \exp \left\{ \Omega_i^T(\phi) + \alpha_1 m_i(t) + \alpha_2 m_i(t') \right\}, \quad (4)$$

where  $m_i(t') = (d/dt)(m_i(t))$ . Additionally, we fitted the association structure to allow the whole history of the blood glucose trajectory to be associated with the risk of the preterm birth. Then, the event model takes the form

$$h_i[t] = lt^{l-1} \exp \left\{ \Omega_i^T(\phi) + \alpha \int_0^t m_i(s) ds \right\}. \quad (5)$$

For any particular time point,  $\alpha$  in (4) measures the strength of association between the risk of PB at time  $t$  and the area under the longitudinal trajectory of the blood glucose level up to the same  $t$ . We tested all three association structures (value, slope, and area under the curve) and selected the best model based on the deviance information criteria (DIC) [28]. The deviance information criteria measure balances, the fit of a model to the data with its complexity. A smaller value of DIC indicates the preferred model.

We compared the joint model with the time-dependent Cox model where we fitted the preterm birth event model with a Cox model that included glucose measurements as a

regular time dependent covariate (ignoring measurement error), i.e.,

$$h_i[t] = t^{l-1} \exp \left\{ \gamma_1(\text{glucose})_{ij} + \gamma_2(\text{gestational day})_{ij} + \gamma_3(\text{glucose})_{ij} * (\text{gestational day})_{ij} + \gamma_4(\text{age})_i + \gamma_5(\text{age at diabetes onset})_i + \gamma_6(\text{presence of nephropathy})_i \right\}. \quad (6)$$

Model comparison between the time-dependent Cox model and the joint model was conducted by DIC. All data analyses were conducted using SAS v9.4.4 (Cary, NC) and JMbayes package in RStudio (2015), where the time-dependent Cox model was estimated using a frequentist procedure, and the joint models were implemented using the approach available in JMbayes. We used readily available software to estimate the model parameters. The default prior distribution was used for JMbayes to estimate all joint model parameters. We preferred to use JMbayes instead of the JM software package, since the former is actively maintained and provides straightforward implementation of dynamic prediction [12]. Estimates from each model are reported with 95% confidence or credible intervals (denoted as 95% CI), depending upon whether the model was implemented with the frequentist or Bayesian approach.

**2.3. Individual Dynamic Prediction.** In this paper, we used Rizopoulos-proposed approach [25] to dynamically update the risk of preterm birth based on the updated glucose trajectory for an individual subject. A key feature of these dynamic prediction frameworks is that the predictive measures for the risk of preterm birth can be dynamically updated as additional longitudinal measurements of the blood glucose level become available for the target subjects, providing instantaneous risk assessment.

### 3. Results

Descriptive statistics are presented in Table 1. A total of 114 women with T1DM with 16,480 glucose readings were included in this analysis. The median (IQR) number of glucose recordings was 157 (6-211); 32 (28%) of the women had preterm delivery. Median age of entry was 27 (22-31) years; median age of diabetes onset was 12 (8-17) years; median baseline BMI of the cohort was 23 (21-25) kg/m<sup>2</sup>. A total of 32 (28%) were diagnosed with preeclampsia, 14 (12%) with chronic hypertension, 40 (35%) with white classification (B/C vs ≥D), 22(19%) with nephropathy, and 16 (14%) with retinopathy.

In order to allow for flexibility in the nonlinear blood glucose profiles, we included a natural cubic splines in both the fixed and random effects part of the mixed effects model. A range of knots was used to fit the longitudinal glucose profiles. Based on Bayesian information criteria, a total of 22 knots was selected to fit individual profiles. Figure 1 corresponds to the overall model fit for the glycemic profiles and 95% CI for the period of 50-260 gestational days. Mother's age at last menstrual period ( $\beta_{\text{age}} = 1.29$ , 95% CI 1.10, 1.72)

TABLE 1: Patient characteristics ( $n = 114$ ; 16,480 glucose readings over gestation 50-260 days).

	Median (IQR)	$f$ (%)
Mother's age at LMP	27 (22-31)	
Age at diabetes onset	12 (8-17)	
Baseline BMI	23 (21-25)	
Number of glucose recordings	157 (6-211)	
PB		32 (28)
Preeclampsia		32 (28)
White classification		
B, C		40 (35)
≥D		74 (65)
Presence of nephropathy		22 (19)
Presence of hypertension		14 (12)
Presence of retinopathy		16 (14)

IQR: Interquartile range;  $f$ : Frequency; BMI: body mass index (kg/m<sup>2</sup>).

and age at diabetes onset ( $\beta_{\text{ageatdiabetesonset}} = 0.84$ , 95% CI 0.62, 0.98) were statistically significant for the longitudinal submodel, reflecting that older women tended to have a higher mean blood glucose level and those with later diabetes onset tended to have a lower mean blood glucose level (Table 2). The presence of nephropathy was statistically significant for the event submodel, reflecting that the probability of preterm birth significantly increases with the presence of nephropathy among women with T1DM ( $RR_{\text{neph}} = 2.29$ , 95% CI 1.05, 4.48). Based on deviance information criteria, the joint model provided better model fit in comparison to the time-dependent Cox model ( $DIC_{\text{JM}} = 19895$ ;  $DIC_{\text{TDCM}} = 19932$ ). Unlike the joint model, mother's age and age at diabetes onset were not significant in the time-dependent Cox model. The presence of nephropathy was statistically significant in the time-dependent Cox model ( $RR_{\text{neph}} = 2.21$ , 95% CI 1.01, 3.32).

Based on deviance information criteria, the association link with cumulative glucose monitoring provided the best fit ( $DIC_{\alpha_{\text{value}}} = 19,990$ ;  $DIC_{\alpha_{\text{value plus slope}}} = 19,905$ ;  $DIC_{\alpha_{\text{AUC}}} = 19,895$ ) compared to the value or rate of the glucose level at time  $t$  (Table 3). The cumulative association (area under the curve) indicated that the cumulative effects of the glucose monitoring outcome up to time point  $t$  had the strongest association ( $\alpha_{\text{AUC}} = -0.01$ ,  $p < 0.01$ ) between the risk of preterm birth and maternal glycemic profile.

The assumption for dynamic prediction is that the woman was event-free (i.e., did not give birth) up to the time point of the last glucose reading available. Using the Markov chain Monte Carlo algorithm, the estimates of the risk of preterm birth were computed for a new subject knowing her glucose readings up to a given gestation day. Figure 2 depicts the change in the risk of preterm birth of a woman with an increasing number of glucose readings over time. For glucose readings up to 107 gestational days, the woman had roughly a 25% risk of having premature delivery. The risk decreases by 5% with her available glucose readings out to 212 gestational days in pregnancy.

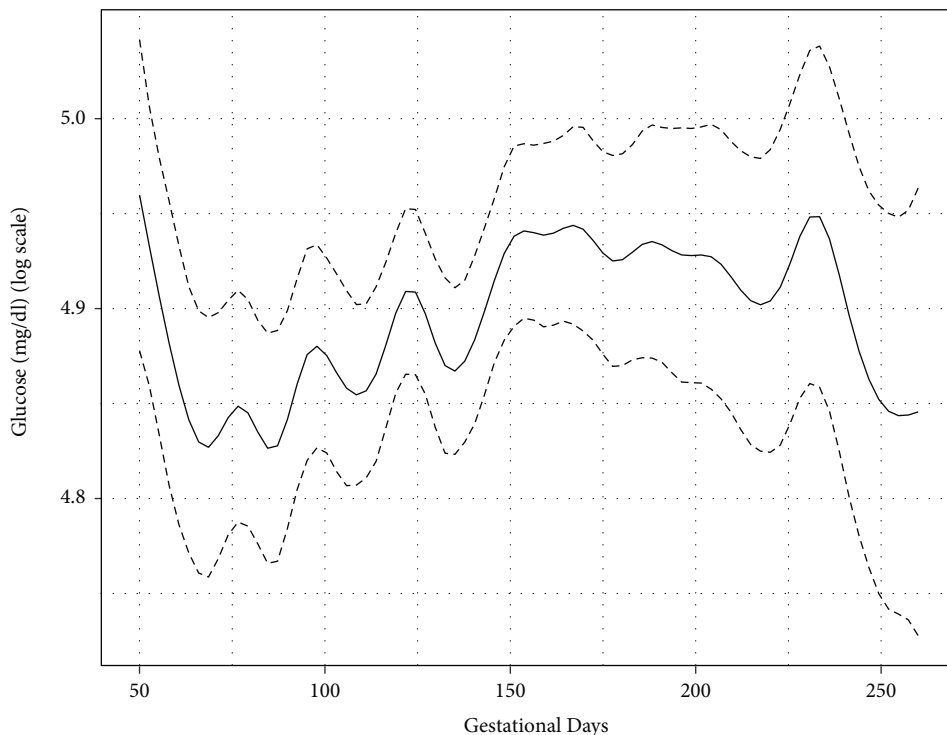


FIGURE 1: Fitted glucose model (on log scale) with natural spline over 50-260 gestational days.

TABLE 2: Parameter estimates and model fit statistics from the joint model and the time-dependent Cox model.

	Time-dependent Cox model			LSM EST (CI)	Joint model ESM	
	EST (CI)	OR (CI)	M-Val		OR (CI)	M-Val
Mother’s age at LMP	1.02 (0.95, 1.10)			1.29 (1.10, 1.72)*		
Age at diabetes onset	0.97 (0.92, 1.03)			0.84 (0.62, 0.98)*		
Presence of nephropathy		2.21 (1.01, 3.32)*			2.29 (1.05, 4.48)*	
Association						-0.01*
DIC			19,932			19,895
-2Log-likelihood			18,464			19,434

\* $p < 0.05$ ; CI refers to confidence and credible interval for the time-dependent Cox model and the joint model, respectively (see Methods); LSM: longitudinal submodel; ESM: event submodel; EST: estimate; OR: odds ratio, M-Val: model value.

TABLE 3: Value, value plus slope, and area under the curve association links and model fit statistics for the joint models.

Association link	DIC	pD
Value	19990	2182
Value and slope	19905	2143
Area under the curve	19895	2130

DIC: deviance information criteria; pD: effective number of parameters. Lower values indicate better model fit.

#### 4. Discussion

In this study, we demonstrated that the characterizing and estimating rate of maternal glucose sampling profile over the course of the entire pregnancy is an important task to predict preterm birth outcome. Recent studies have shown

that the rate of both spontaneous and indicated preterm birth is increased in women with T1DM [29, 30]. Preterm babies go on to have a higher risk of increased morbidity, mortality, and poor neurological outcome [31]. In mothers with T1DM, poor glycemic control was associated with both spontaneous and indicated preterm birth [32]. As we have shown in our previously published work [8], nonlinear curvature of blood glucose monitoring data makes it analytically challenging. In recent years, the joint model has been shown to be an efficient approach to simultaneously model longitudinal data and an observed outcome, often yielding more accurate, informative prediction than traditional models [12]. In this paper, we describe a joint model approach aimed at analyzing long sequences of longitudinal and time-to-event data and used it to simultaneously characterize the glycemic profile for women with T1DM and assess risk of preterm birth. On the population level, we identified demographic

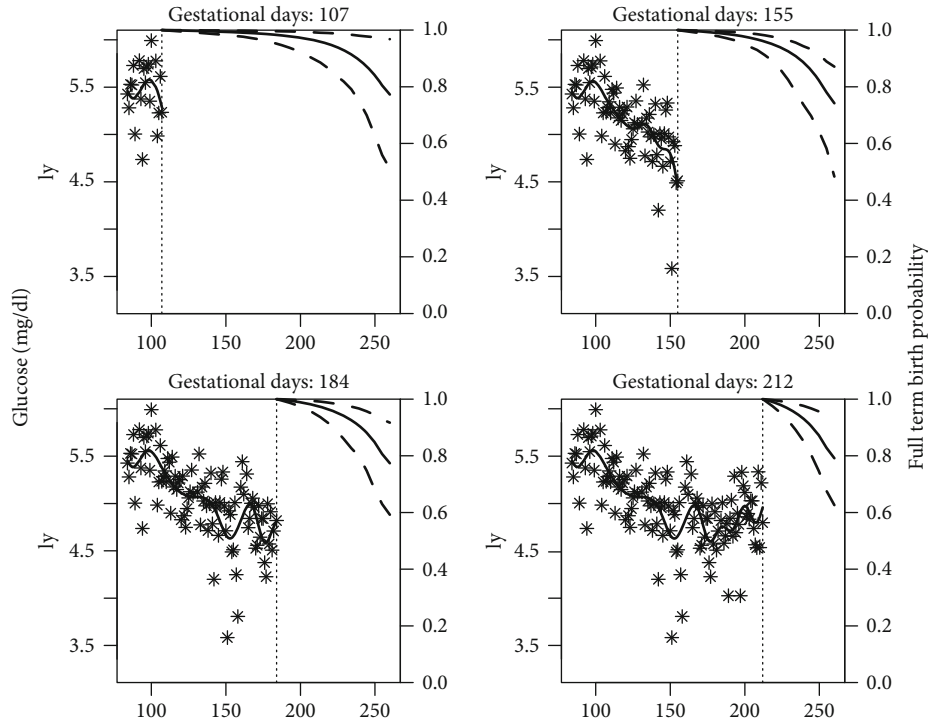


FIGURE 2: Individual dynamic prediction of risk for preterm birth with updating blood glucose readings (on log scale).

characteristics associated with longitudinal and event outcomes. When comparing the joint model approach to the commonly used time-dependent Cox model, we found that the former was a better fit based on the model fit statistics and characterized association between the risk of preterm birth and glucose measurements over the entire pregnancy period to provide additional insights compared to the time-dependent Cox model. While maternal age at last menstrual period and age at diabetes onset were not significant in the time-dependent Cox model, it was significant in the joint model providing additional information regarding the relationship between maternal glycemic profile over the course of entire pregnancy and mother's age at last menstrual period and age at diabetes onset. We found that older women tended to have higher mean blood glucose, and those with later diabetes onset tended to have lower mean blood glucose. The presence of nephropathy was statistically significant in both the joint model and the time-dependent Cox model.

We tested various association structures to share the parameters between the longitudinal and event submodel within the joint modeling framework. Based on deviance information criteria, the association structure based on area under the curve of the glycemic profile provided the best fit for the model in comparison to the value and value and rate association structures. An alternative criterion for Bayesian joint models is the logarithm of the pseudo likelihood (LPML) [33], which may be computed separately for the submodels using a Monte Carlo approach; however, readily available software does not include either version of the LPML. BIC may be used to rank joint models but, as the authors of the LPML approach describe, it may not be considered in submodel-specific assessments without proper

decomposition. The cumulative association structure, as defined by the area under the curve, is clinically more appealing as it includes the cumulative effects of the glucose monitoring outcome up to a certain time point to determine the risk of preterm birth of a subject. Further, our approach, which includes a linear mixed effects submodel, accounts for missing glucose data. This is not possible in summary calculations, which are prone to measurement error [34].

A novel use of joint models, leveraged in this study, is to obtain dynamic personalized prediction of future longitudinal outcome trajectories and risks of an event at any time, given the subject-specific outcome profiles up to the time of prediction. Rizopoulos proposed a Monte Carlo approach [25] to estimate risk of a target event and illustrated how it can be dynamically updated. We followed this approach to show how dynamic prediction for the risk of preterm birth can be obtained as a function of maternal glycemic profile over the course of an entire pregnancy for a woman in this T1DM cohort. Based on the fitted joint model, we derived the prediction for a new subject from the same population that provided a set of longitudinal glucose measurements up to a specified time. As an example, we considered one maternal glycemic profile and, based on a Monte Carlo procedure, showed the mean and median estimates over the Monte Carlo samples along with the 95% pointwise credible intervals. We calculated the conditional full-term birth (i.e., event-free) probabilities for an individual patient, starting from her first glucose measurements and adding glucose at each time point as an extra measurement.

There are some limitations to the current study. We conducted variable selection based on a one-at-a-time model selection process. Although this approach is suitable for

prediction modeling, it cannot be used to infer causality; therefore, association estimates from this study must be viewed cautiously [35]. With respect to missing data in maternal glycemic profiles, our approach assumes that data are missing at random [36] in the sense that missingness can be explained by observed information. By using the existing R package, JMbayes, we modeled observed birth outcome as a survival process, with no censoring in the data. Our further research is aimed at rectifying this problem by including the submodel as a binary process. The joint model [12, 13, 37, 38] approach proposed in this study may be used to model continuous glucose monitoring data, thus incorporating the cutting-edge technology in diabetes care. Moreover, future work should aim to separately model the irregular glucose sampling pattern associated with the blood glucose level for women with T1DM to predict the risk of preterm birth. Another interesting future work area will be to look into the effects of glycohemoglobin along with glucose measurements to predict the risk of preterm birth using the joint modeling technique.

## 5. Conclusions

We have developed a joint model approach that, once applied to motivating data, provided more realistic estimates of the maternal glycemic profile over the entire pregnancy and enabled individualized assessment of preterm birth risk in women with T1DM. Through a Bayesian Markov chain Monte Carlo algorithm implemented in open-source software, this model could be used to aid clinicians in estimating the risk of preterm delivery sooner than later based on the glucose trajectory and other baseline covariates, potentially allowing for intervention strategies to be applied. We demonstrated that overall model fit improved with our joint model, compared to the conventional time-dependent Cox model. Some important covariates, mother's age at LMP and age of diabetes onset, were significant in the joint model unlike using the conventional approach, the time-dependent Cox model. Accurately characterizing the association between blood glucose level and preterm birth may serve as a prognostic aid for clinical decision making allowing for research focused on personalized, preemptive treatment and monitoring of pregnancies impacted by T1DM.

## Data Availability

The data used in the present study to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study is supported by R01 DK109956; Level and timing of diabetic hyperglycemia in utero: the transgenerational effect on adult morbidity (IRB 2017-1236).

## References

- [1] J. K. Kirk and J. Stegner, "Self-monitoring of blood glucose : practical aspects," *Journal of Diabetes Science and Technology*, vol. 4, no. 2, pp. 435–439, 2010.
- [2] O. Langer and D. L. Conway, "Level of glycemia and perinatal outcome in pregestational diabetes," *The Journal of Maternal-Fetal Medicine*, vol. 9, no. 1, pp. 35–41, 2000.
- [3] D. Deiss, J. Bolinder, J. P. Riveline et al., "Improved glycemic control in poorly controlled patients with type 1 diabetes using real-time continuous glucose monitoring," *Diabetes Care*, vol. 29, no. 12, pp. 2730–2732, 2006.
- [4] O. Kovilam, J. Khoury, M. Miodovnik, M. Chames, J. Spinnato, and B. Sibai, "Spontaneous preterm delivery in the type 1 diabetic pregnancy: the role of glycemic control," *Journal of Maternal Fetal Neonatal Medicine*, vol. 11, no. 4, pp. 245–248, 2002.
- [5] J. Lepercq, J. Coste, A. Theau, D. Dubois-Laforgue, and J. Timsit, "Factors associated with preterm delivery in women with type 1 Diabetes: A cohort study," *Diabetes Care*, vol. 27, no. 12, pp. 2824–2828, 2004.
- [6] E. G. Deryabina, G. V. Yakornova, L. A. Pestrvaeva, and N. D. Sandyreva, "Perinatal outcome in pregnancies complicated with gestational diabetes mellitus and very preterm birth: case-control study," *Gynecological Endocrinology*, vol. 32, Supplement 2, pp. 52–55, 2016.
- [7] S. F. Lin, C. F. Kuo, M. J. Chiou, and S. H. Chang, "Maternal and fetal outcomes of pregnant women with type 1 diabetes, a national population study," *Oncotarget*, vol. 8, no. 46, pp. 80679–80687, 2017.
- [8] R. Gupta, J. Khoury, M. Altaye, L. Dolan, and R. D. Szczesniak, "Glycemic excursions in type 1 diabetes in pregnancy: a semi-parametric statistical approach to identify sensitive time points during gestation," *Journal of Diabetes Research*, vol. 2017, Article ID 2852913, 7 pages, 2017.
- [9] P. K. Andersen and R. D. Gill, "Cox's regression model for counting processes: a large sample study," *The Annals of Statistics*, vol. 10, no. 4, pp. 1100–1120, 1982.
- [10] P. RL, "Covariate measurement errors and parameter estimation in a failure time regression model," *Biometrika*, vol. 69, no. 2, pp. 331–342, 1982.
- [11] J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, New York, 2nd edition, 2002.
- [12] D. Rizopoulos, *Joint Models for Longitudinal and Time-to-Event Data, with Applications in R*, Chapman & Hall/CRC, Boca Raton, 2012.
- [13] D. Rizopoulos and E. Lesaffre, "Introduction to the special issue on joint modelling techniques," *Statistical Methods in Medical Research*, vol. 23, no. 1, pp. 3–10, 2014.
- [14] M. S. Wulfsohn and A. A. Tsiatis, "A joint model for survival and longitudinal data measured with error," *Biometrics*, vol. 53, no. 1, pp. 330–339, 1997.
- [15] A. A. Tsiatis and M. Davidian, "Joint modeling of longitudinal and time-to-event data: an overview," *Statistica Sinica*, vol. 14, pp. 809–834, 2004.
- [16] V. De Gruttola and X. M. Tu, "Modelling progression of CD-4 lymphocyte count and its relationship to survival time," *Biometrics*, vol. 50, no. 4, pp. 1003–1014, 1994.
- [17] Y. Wang and J. M. G. Taylor, "Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome," *JAMA*, vol. 96, no. 455, pp. 895–905, 2001.

- [18] R. Henderson, P. Diggle, and A. Dobson, "Joint modelling of longitudinal measurements and event time data," *Biostatistics*, vol. 1, no. 4, pp. 465–480, 2000.
- [19] E. R. Brown and J. G. Ibrahim, "A Bayesian semiparametric joint hierarchical model for longitudinal and survival data," *Biometrics*, vol. 59, no. 2, pp. 221–228, 2003.
- [20] D. Rizopoulos and P. Ghosh, "A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event," *Statistics in Medicine*, vol. 30, no. 12, pp. 1366–1380, 2011.
- [21] L. Wu, W. Liu, G. Y. Yi, and Y. Huang, "Analysis of longitudinal and survival data: joint modeling, inference methods, and issues," *Journal of Probability and Statistics*, vol. 2012, Article ID 640153, 17 pages, 2012.
- [22] W. Ye, X. H. Lin, and J. M. G. Taylor, "Semiparametric modeling of longitudinal measurements and time-to-event data – a two-stage regression calibration approach," *Biometrics*, vol. 64, no. 4, pp. 1238–1246, 2008.
- [23] X. Huang, G. Li, R. M. Elashoff, and J. Pan, "A general joint model for longitudinal measurements and competing risks survival data with heterogeneous random effects," *Lifetime Data Analysis*, vol. 17, no. 1, pp. 80–100, 2011.
- [24] M. Yu, N. Law, J. Taylor, and H. Sandler, "Joint longitudinal-survival-cure models and their application to prostate cancer," *Statistica Sinica*, vol. 14, pp. 835–862, 2004.
- [25] D. Rizopoulos, "Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data," *Biometrics*, vol. 67, no. 3, pp. 819–829, 2011.
- [26] S. S. McElvy, M. Miodovnik, B. Rosenn et al., "A focused pre-conceptional and early pregnancy program in women with type 1 diabetes reduces perinatal mortality and malformation rates to general population levels," *The Journal of Maternal - Fetal Medicine*, vol. 9, no. 1, pp. 14–20, 2000.
- [27] A. Gelfand and D. Dey, "Bayesian model choice: asymptotics and exact calculations. (series B)," *Journal of the Royal Statistical Society*, vol. 56, pp. 501–514, 1994.
- [28] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde, "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 64, no. 4, pp. 583–639, 2002.
- [29] S. Li, M. Zhang, H. Tian, Z. Liu, X. Yin, and B. Xi, "Preterm birth and risk of type 1 and type 2 diabetes: systematic review and meta-analysis," *Obesity Reviews*, vol. 15, no. 10, pp. 804–811, 2014.
- [30] I. Eidem, S. Vangen, K. F. Hanssen et al., "Perinatal and infant mortality in term and preterm births among women with type 1 diabetes," *Diabetologia*, vol. 54, no. 11, pp. 2771–2778, 2011.
- [31] J. F. Ludvigsson, M. Neovius, J. Söderling et al., "Maternal Glycemic Control in Type 1 Diabetes and the Risk for Preterm Birth: A Population-Based Cohort Study," *Annals of Internal Medicine*, vol. 170, no. 10, pp. 691–701, 2019.
- [32] A. Z. Feldman and F. M. Brown, "Management of type 1 diabetes in pregnancy," *Current Diabetes Reports*, vol. 16, no. 8, p. 76, 2016.
- [33] D. Zhang, M. H. Chen, J. G. Ibrahim, M. E. Boye, and W. Shen, "Bayesian model assessment in joint modeling of longitudinal and survival data with applications to cancer clinical trials," *Journal of Computational and Graphical Statistics*, vol. 26, no. 1, pp. 121–133, 2017.
- [34] G. M. Fitzmaurice, N. M. Laird, and J. H. Ware, *Applied Longitudinal Analysis*, Wiley, New York, 2nd edition, 2011.
- [35] D. J. Lederer, S. C. Bell, R. D. Branson et al., "Control of confounding and reporting of results in causal inference studies. Guidance for authors from editors of respiratory, sleep, and critical care journals," *Annals of the American Thoracic Society*, vol. 16, no. 1, pp. 22–28, 2019.
- [36] G. Molenberghs, G. Fitzmaurice, M. G. Kenward, A. Tsiatis, and G. Verbeke, *Handbook of Missing Data Methodology*, Chapman & Hall, 2015.
- [37] G. L. Hickey, P. Philipson, A. Jorgensen, and R. Kolamunnage-Dona, "Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues," *BMC Medical Research Methodology*, vol. 16, no. 1, p. 117, 2016.
- [38] G. L. Hickey, P. Philipson, A. Jorgensen, and R. Kolamunnage-Dona, "Joint models of longitudinal and time-to-event data with more than one event time outcome: a review," *International Journal of Biostatistics*, vol. 14, no. 1, 2018.