Received: 14 January 2019 | Revised: 5 September 2019 | Accepted: 6 September 2019

DOI: 10.1002/bimj.201900044

Biometrical Journal

RESEARCH PAPER

A tutorial on dynamic risk prediction of a binary outcome based on a longitudinal biomarker

Rana Dandis[1] | **Steven Teerenstra**[1] | **Leon Massuger**[2] | **Fred Sweep**[3] | **Yalck Eysbouts**[2] | **Joanna IntHout**[1]

[1] Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

[2] Department of Obstetrics and Gynecology, Radboud University Medical Center, Nijmegen, The Netherlands

[3] Department of Laboratory Medicine, Radboud University Medical Center, Nijmegen, The Netherlands

Correspondence
Rana Dandis, Radboud Institute for Health Sciences, Radboud University Medical Center, Geert Grooteplein noord 21, 6500 HB Nijmegen, The Netherlands.
Email: rana.dandis@radboudumc.nl

Abstract

Dynamic risk predictions based on all available information are useful in timely identification of high-risk patients. However, in contrast with time to event outcomes, there is still a lack of studies that clearly demonstrate how to obtain and update predictions for a future binary outcome using a repeatedly measured biomarker. The aim of this study is to give an illustrative overview of four approaches to obtain such predictions: likelihood based two-stage method (2SMLE), likelihood based joint model (JMMLE), Bayesian two-stage method (2SB), and Bayesian joint model (JMB). We applied the approaches to provide weekly updated predictions of post–molar gestational trophoblastic neoplasia (GTN) based on age and repeated measurements of human chorionic gonadotropin (hCG). Discrimination and calibration measures were used to compare the accuracy of the weekly predictions. Internal validation of the models was conducted using bootstrapping. The four approaches resulted in the same predictive and discriminative performance in predicting GTN. A simulation study showed that the joint models outperform the two-stage methods when we increase the within- and the between-patients variability of the biomarker. The applicability of these models to produce dynamic predictions has been illustrated through a comprehensive explanation and accompanying syntax (R and SAS$^{®}$).

KEYWORDS

binary outcome, dynamic prediction, joint model, longitudinal biomarker, two-stage method

1 | INTRODUCTION

Often, patients' biomarkers are repeatedly measured over time during follow-up visits. In this setting, a key quantity of interest is the prediction of a future outcome based on all available information known up to a certain point of time. These predictions are updated as new extra measurements become available and can be useful in timely identification of high-risk patients, who may benefit from early intervention or treatment.

The use of longitudinal measurements to obtain updated predictions of the risk of a future binary outcome cannot be done efficiently with a classical model like multiple logistic regression, because this is not particularly suited to directly model changes of the biomarker over time. Moreover, the longitudinal biomarkers are usually measured with error, which would tend to attenuate the coefficient of the regression model reflecting the relation between the risk of the future outcome and the biomarker, leading to underestimation of the association between the biomarker and the risk of the outcome. To address this, one could consider including the repeated measurements of the biomarker by using summary measures that capture their change over time. Ideally for this situation, the summary measures could be latent variables (e.g., subject-specific random intercept and slope effects) that summarize the evolution of the longitudinal data over time through a mixed effects model. Then, these summaries can be linked to the subject-specific risk of developing the binary outcome using logistic regression, either in separate steps through a two-stage approach (Wang, Wang, & Wang, 2000), or by using a joint modeling approach where simultaneous estimation of parameters from the longitudinal biomarker and the binary outcome is done (Horrocks & van Den Heuvel, 2009). The early development of the two-stage method and joint modeling was mainly focused on the relation between a longitudinal predictor and a survival outcome (Ibrahim, Chu, & Chen, 2010; Rizopoulos, 2012; Wulfsohn & Tsiatis, 1997). However, in many applications, the exact timing of the event is either unknown or not informative. For such cases, analogous models for a longitudinal predictor with a binary outcome are needed.

In this paper, we propose a framework of four possible approaches to predict the risk of a future binary outcome based on a repeatedly measured predictor. We consider the two-stage method and the joint modeling approach, using two different estimation methods: maximum likelihood and Bayesian. Moreover, using the resulting models, we show how to obtain dynamically updated subject-specific risk predictions for new patients at each time additional measurements of the longitudinal biomarker are recorded. We first illustrate the four modeling approaches using a real example data set, and then we show how the performance of these approaches differs in response to changes in the within- and the between-subject variability using simulations. We explain the approaches with theory and programming syntax (R and SAS®) (Inc 2015; Team, 2017), and show that the discussed approaches together with the accompanied syntax are a useful toolbox to obtain updated predictions based on data recorded over time.

2 | EXAMPLE DATA SET

Gestational trophoblastic diseases (GTD) is a group of pregnancy-related diseases including hydatidiform moles. Although curable in most cases, progression to invasive and life-threatening disease occurs in 10–15% of cases (Stevens et al., 2015). Then, the removal of the trophoblastic tissue by suction curettage (evacuation) is often performed as it is the recommended treatment of choice after diagnosis of GTD. After this evacuation procedure, patients are at risk for post–molar gestational trophoblastic neoplasia (GTN), which is characterized by persistently elevated serum levels of human chorionic gonadotropin (hCG) (Cavaliere, Ermito, Dinatale, & Pedata, 2009).

Several studies discussed the challenge of predicting post–molar GTN based on either single hCG measurements or a summary measure like the decline rate that captures the change in hCG over time (Aminimoghaddam et al., 2014; Kim et al., 2012; Mousavi, Karimi, Modarres Gilani, Akhavan, & Rezayof, 2014). Eysbouts et al. developed a practical and easy tool to obtain subject-specific risk predictions of post–molar GTN for each of the first four weeks after evacuation, based on serum hCG (Eysbouts, Massuger, Ottevanger, IntHout, & Sweep, 2017). Recently, Khosravirad et al. investigated the use of subject-specific random effects as summary measures to make the best use of the weekly collected hCG measurements to predict post–molar GTN using a two-stage model; they showed that using hCG levels for the last two weeks is as good as using hCG levels from all the four follow-up weeks (Khosravirad, Zayeri et al. 2017).

Motivated by the importance of early detection of GTN in women with molar pregnancy, we use the data obtained from the Dutch Central Registry for hydatidiform moles at the Radboud University Medical Center in Nijmegen (Radboudumc). All patients with available serum hCG data obtained after mole evacuation surgery were included. The main outcome of our study is the presence of GTN, which is a binary outcome. The GTN status was determined only after following the patients up to a maximum of seven weeks. Since the exact time of GTN development is unknown, a model for a binary outcome is needed. There were 299 women in the uneventful group (GTN = 0), and 140 patients with a confirmed diagnosis of post molar GTN in the persistent trophoblastic disease (GTN = 1) group. Serum hCG levels (ng/mL) taken between two and seven weeks after evacuation were retrospectively evaluated for all women. A total of 1674 serum hCG measurements were available, ranging between one and six measurements and with a median of four measurements per woman. The hCG levels were 10 log-transformed and assigned per “week since evacuation.”

Plotting subject-specific data is always recommended before conducting any data analyses. Figure 1 represents the longitudinal profiles of the 10 log-transformed hCG measurements over time using a spaghetti plot, which is useful for graphically

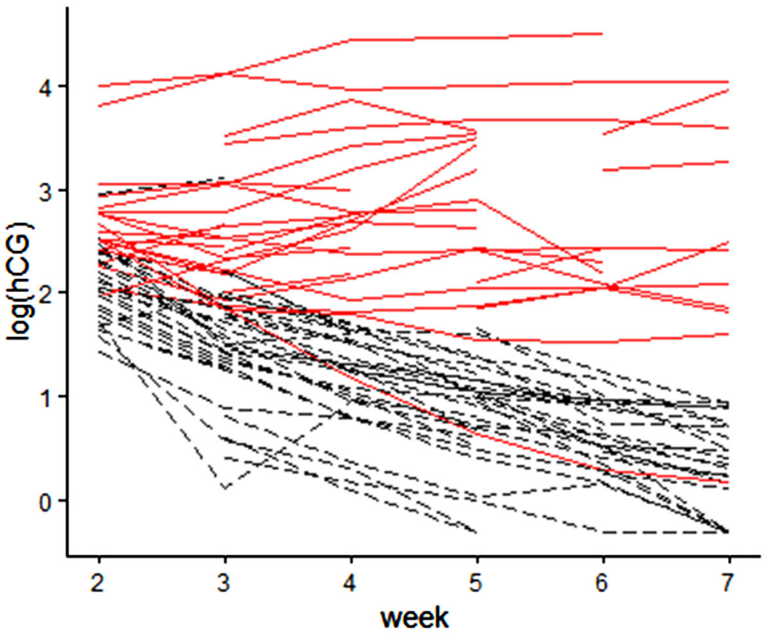


FIGURE 1 10 log-transformed hCG profiles for 100 randomly selected women in the first seven weeks after mole evacuation (solid red lines = GTN patients, dashed black lines = uneventful women)

displaying the change of the longitudinal data over time. In general, the hCG levels in the group of patients who develop GTN are higher and more stable compared to the hCG levels in the uneventful group, which start at lower levels and decrease with time. Women with post–molar GTN presented at a slightly older age than those of the uneventful group, with mean ages of 31.2 (*SD* = 7.0) and 30.0 (*SD* = 6.9) years, respectively.

3 | STATISTICAL FRAMEWORK

In this section, we discuss a conceptual framework, presented in Figure 2, for providing updated risk prediction of a binary outcome based on longitudinal measurements, using the following four modeling approaches:

1. maximum likelihood–based two-stage method (2SMLE),
2. maximum likelihood–based joint model (JMMLE),
3. Bayesian two-stage method (2SB), and
4. Bayesian joint model (JMB).

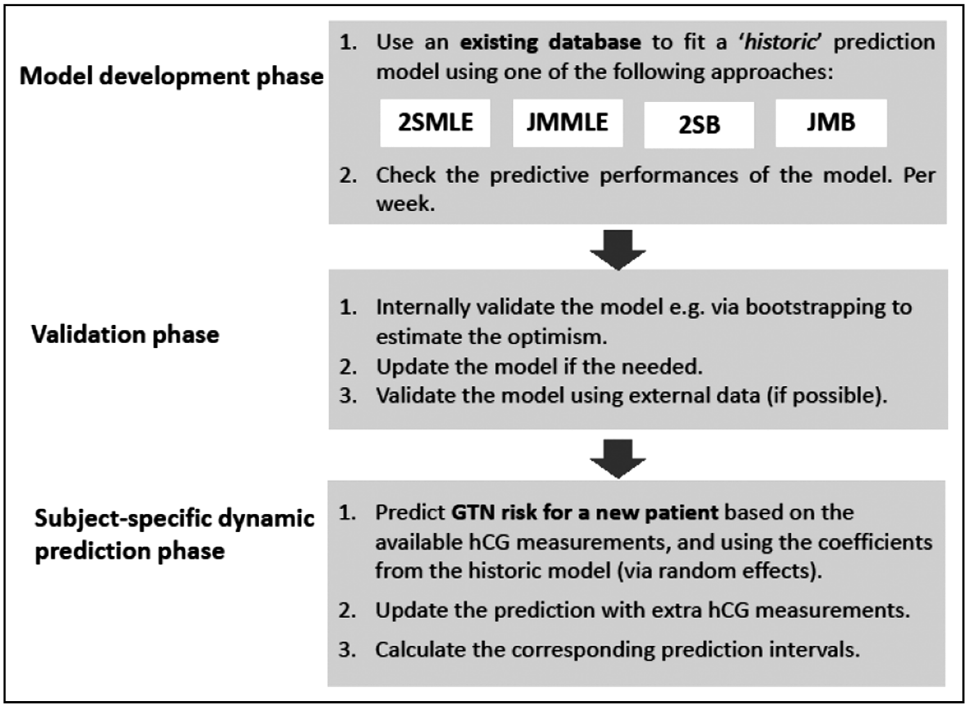
We applied the four approaches using the GTN database, from which we excluded two patients, whose data were used later (Section 4.3) to illustrate the dynamic prediction phase for new patients. We used R software version 3.3.3 (Team, 2017), SAS® version 9.4 (Inc, 2015), and JAGS 4.3.0, and we provided the detailed syntax to implement the above framework, see the supplementary material on the journal's web page (http://onlinelibrary.wiley.com/doi/10.1002/bimj.201900044/suppinfo).

3.1 | Development of the historic model

3.1.1 | The two-stage approach

The two-stage approach was proposed by Tsiatis et al. to analyze a longitudinal predictor in combination with survival data (Tsiatis, Degruttola, & Wulfsohn, 1995). We apply the same concept, but for a binary outcome, where in the first stage, the whole evolutions of the repeated biomarker measurements are summarized by random effects obtained by fitting a linear mixed model, and in the second stage, the resulting random effects are used as covariates in a logistic regression model to predict the risk of the future outcome (GTN).

A key step in the model development is specifying the function for the fixed and the random effects. In our example, the most plausible model to describe the change of the log(hCG) over time is a random intercept and random slope model to capture both the variation in the starting levels of hCG and the linear change of these measurements over time (week 2 until week 7) (see Figure 1). However, the random intercept and random slope model may not always be the best fit for the data, for example, the



Model development phase

1. Use an **existing database** to fit a *'historic'* prediction model using one of the following approaches:

2SMLE JMMLE 2SB JMB

2. Check the predictive performances of the model. Per week.

Validation phase

1. Internally validate the model e.g. via bootstrapping to estimate the optimism.
2. Update the model if the needed.
3. Validate the model using external data (if possible).

Subject-specific dynamic prediction phase

1. Predict **GTN risk for a new patient** based on the available hCG measurements, and using the coefficients from the historic model (via random effects).
2. Update the prediction with extra hCG measurements.
3. Calculate the corresponding prediction intervals.

FIGURE 2 The conceptual framework for dynamic prediction of a binary outcome based on longitudinal biomarker measurements

change of the biomarker could be quadratic in time, and in this case to better capture the shape of the subject-specific longitudinal trajectories, we need to include extra random-effect terms. This could be achieved using, for instance, polynomials or regression splines.

The inclusion of patients' baseline characteristics, such as age at evacuation, along the hCG repeated measurements may improve GTN prediction. Age may have an effect on both the evolution of hCG over time as well as the development of GTN. Hence, and also for illustrative purposes, we included age in both model stages. In other situations, it might be sufficient to include additional predictors like age only in one of the stages, for example, only in the prediction stage if the predictor is only relevant for the binary outcome but not for the longitudinal profile of the biomarker.

Let $\log(\text{hCG})_{it}$ represent the 10 log-transformed hCG longitudinal measurements for patient i, $i = 1, \ldots, 439$, at week t, $t = 2, \ldots, 7$. The model for the first stage can be written as follows:

$$\log(\text{hCG})_{it} = \beta_0 + b_{0i} + \left(\beta_1 + b_{1i}\right) t + \beta_2 \text{AGE}_i + \epsilon_{it}, \tag{1}$$

where β_0, β_1, and β_2 are unknown fixed effect parameters, b_{0i} and b_{1i} are unknown patient-specific random intercept and slope, respectively, which are assumed to have a bivariate normal distribution with mean zero and covariance matrix $\boldsymbol{D}$; AGE_i is the patient age at the time of evacuation; and ϵ_{it} is the residual error for patient i at week t with a normal distribution $N(0, {\sigma_e}^2)$, which is assumed to be independent of the random effects. Initially, we included both linear and quadratic effects in the linear mixed model to describe the change of hCG over time. The model with quadratic curves resulted in a better fit to the data, with a lower AIC (Akaike information criterion) when compared to the model with linear trajectories. However, both models resulted in the same prediction accuracy when used later to predict GTN, so we decided to use the model that assumes linear trajectories on the grounds of parsimony.

In the second stage, patient age as well as the point estimates of the subject-specific random effects, b_{0i} and b_{1i} (see Section 3.1.3 and Figure 3 for more information on how the random effects are estimated), from Stage 1 were used as predictors in a logistic regression model with the status of GTN as the outcome:

$$\text{logit}\left(\text{P}\left(\text{GTN}_i = 1\right)\right) = \alpha_0 + \alpha_1 \hat{b}_{0i} + \alpha_2 \hat{b}_{1i} + \alpha_3 \text{AGE}_i, \tag{2}$$

where GTN_i reflects the GTN status of patient i, and $\alpha = [\alpha_0, \alpha_1, \alpha_2, \alpha_3]$ is the vector of the logistic regression coefficients. The coefficients α_1 and α_2 reflect the strength of association between the two models. When this association exists, the use of the longitudinal data improves the predictive ability of the logistic regression model compared to the reduced model using only age as a baseline covariate.

In the above two-stage approach, the logistic model for GTN conditions on the estimated random effects ($\hat{b}_{0i}$ and $\hat{b}_{1i}$) from the mixed model of hCG. However, this approach ignores that these random effects are not exactly known but estimated, which may lead to underestimation as well as imprecision in the regression coefficients. Therefore, a longitudinal submodel with a good fit is needed to produce sufficient random effect estimates that capture the relevant change of hCG over time.

3.1.2 | The joint modeling approach

Joint modeling is a statistical technique that is used to estimate common parameters of two or more models simultaneously (Emmanuel Lesaffre, 2013). Here, we use a joint model to combine the submodels (1) and (2) of the two-stage approach in one model. If we let π_i represent the probability of developing post–molar GTN, that is, $P(GTN_i = 1)$, the likelihood L for the joint model can be written as follows:

$$L = \prod_{i=1}^{n} \int \prod_{t=2}^{7} \varphi\left(\epsilon_{it}\right) \left({\pi_i}^{GTN_i}\left(1-\pi_i\right)^{1-GTN_i}\right) \varphi\left(b_i\right) db_i, \tag{3}$$

where

$$\pi_i = 1/\left(1 + e^{-(\alpha_0+\alpha_1\ b_{0i}+\alpha_2 b_{1i}+\alpha_3 \mathrm{AGE}_i)}\right) \tag{4}$$

$$\epsilon_{it} = \log\left(hCG\right)_{it} - \left(\beta_0 + b_{0i} + \left(\beta_1 + b_{1i}\right) t + \beta_2 \mathrm{AGE}_i\right), \tag{5}$$

$\varphi(\epsilon_{it})$ represents a normal density function with mean 0 and variance ${\sigma_e}^2$. Likewise, $\varphi(b_i)$ a bivariate normal density with mean zero and covariance matrix $\boldsymbol{D}$.

The simultaneous estimation of the joint model parameters avoids the potential problem of the biased estimation in the two-stage approach, because the joint model corrects automatically for the imprecision in the estimates of the covariates of the second model.

3.1.3 | Estimation

The likelihood approach

In the first stage of the 2SMLE, the restricted maximum likelihood (REML) is used to estimate the fixed effects (time, age), and the empirical best linear unbiased prediction (EBLUP) to estimate the subject-specific random effects in the linear mixed model. In the second stage, the estimated random effects plus additional relevant covariates are incorporated as predictors in the logistic regression, which is fitted using the maximum likelihood method. The 2SMLE can be fitted using standard regression tools of statistical software.

Unlike the 2SMLE, the joint model JMMLE requires the manual specification of the joint distribution of the two models, which is used to obtain the likelihood function by the integration over the random effects (b_i) in Equation (3). We used the parameter estimates of the 2SMLE method as initial values for the JMMLE estimation. Alternatively, one could search a grid of initial parameter values. The resulting likelihood is maximized with respect to all its parameters simultaneously. Later, the random effects are obtained with empirical Bayes methodology.

The Bayesian approach

Similar to the 2SMLE, fitting the Bayesian two-stage model (2SB) requires first to estimate the fixed effects (time, age) of the Bayesian linear mixed model as well as the means of the posterior distribution of the random effects, followed by estimating the logistic model parameters in a second stage. This requires specification of the prior distributions for the parameters of the submodels in each stage, that is, $p(\beta)$ in the linear mixed effects model and $p(\alpha)$ in the logistic regression model. Since our prior knowledge is limited, we used proper but vague prior distributions, which are commonly used for the model's location and dispersion parameters. However, in other cases, if prior information is available, informative priors could be used.

Stage I priors: We took normal priors with mean zero and variance 10^2 for the regression parameters β. An inverse gamma prior with shape and scale parameters of 10^{-2} was used for the residual error variance σ_e^2. A multivariate Wishart ($\mathrm{I}_{2\times2}$, 3) distribution was used for the covariance matrix $\boldsymbol{D}$ of the random effects.

Stage II priors: We used normal priors with mean zero and variance 10^2 for the regression parameters α.

The same prior distributions were used to fit the JMB. In this model, the posterior distribution $P(\alpha, \beta|\boldsymbol{y}, \mathrm{X}, \mathrm{Z})$ is proportional to the product $P(\alpha)P(\beta)L(\alpha, \beta|\boldsymbol{y}, \mathrm{X}, \mathrm{Z})$, where $\boldsymbol{y}$ is the vector of the available repeated measurements, and X and Z are the fixed and the random effects design matrices, respectively. We applied the Markov chain Monte Carlo (MCMC) technique, where two chains were initiated with 1000 burn-in iterations and were run for 10,000 iterations.

3.1.4 | The dynamic predictive performance

In this section, we used the following dynamic performance measures to evaluate the dynamic predictive accuracy of the four models (see Section 3.2 for more details on how to obtain the dynamic predictions):

1. the dynamic area under the receiver operator characteristic curve (AUC(t)); a standard measure for estimating the accuracy of a binary classification using a continuous marker;
2. the dynamic mean-squared error of prediction (MSEP(t)), which denotes the average squared difference between the predicted probability and the observed GTN outcome; and
3. the dynamic misclassification error rate (MCER(t)), which is the number of patients who were misclassified (using a cutoff of 50% for the risk of GTN) divided by the total number of patients.

The above dynamic measures are based on the predictions obtained using available longitudinal biomarker measurements until the time of prediction t and the historic models, for example, AUC(t) for the JMB represents the AUC using the historic JMB and the repeated hCG measurements that are known up to week t.

The confidence intervals of the dynamic AUC(t)s and the MSEP(t)s in the likelihood approaches were obtained using bootstrap sampling, where each bootstrap sample was generated by drawing patients with replacement from the original data set. In the Bayesian approaches, the credible intervals of the AUC(t)s and the MSEP(t)s were obtained using the quantiles of the corresponding posterior distributions.

3.1.5 | The internal validation of the dynamic prediction performance

The predictive performances in the above section were assessed using the same data set on which the models were developed, which implies that these are apparent predictive performances and might be optimistic. The most suitable method to validate a model is an external validation, where we apply our fitted models to a new population. However, external validation is not always possible, therefore an unbiased "internal" validation could be applied to evaluate the optimism in the model performance. We used the bootstrap internal validation approach described by Harrell, Lee, and Mark (1996), where the updated predictions at each time point and for each model were validated using the following steps:

1. The historic model was fitted using the original dataset, including all n patients and all time points.
2. The apparent performance measures ($AUC(t)_{app}$, $MSEP(t)_{app}$, and $MCER(t)_{app}$) were obtained using subset from the original data set with the available biomarker measurements up to the corresponding time t.
3. A sample of size n was generated with replacement from the original data set, where patients were drawn by identification number.
4. A new model was fitted using the generated sample in Step 3 and including all n patients and all time points.
5. The performance measures for the model in Step 4 was tested on the data in Step 3 but using only the data up to time t, that is, $AUC(t)_{boot,boot}$, $MSEP(t)_{boot,boot}$, and $MCER(t)_{boot,boot}$.
6. The performance of the model in Step 4 was also evaluated on a subset from the original data set up to time t, that is, $AUC(t)_{boot,orig}$, $MSEP(t)_{boot,orig}$, and $MCER(t)_{boot,orig}$.
7. The optimism in the prediction performance was estimated by subtracting the measures in Step 5 and Step 6, for example, $Optimism(AUC(t)) = AUC(t)_{boot,boot} - AUC(t)_{boot,orig}$.
8. Steps 3 to 7 were repeated 100 times, and the average of the optimism values was calculated.
9. The final bootstrap-corrected updated performance measures were obtained by subtracting the optimism from the apparent measure in Step 2, for example, $AUC(t)_{adj} = AUC(t)_{app} - Optimism(AUC(t))$.

3.2 | Subject-specific dynamic prediction phase

In this phase, based on the fitted historic model, the updated predictions for an individual patient were computed in two steps. First, the random effects $\hat{\boldsymbol{b}}_{it}$, reflecting the hCG profile corrected for age for a new patient i until week t, are predicted from the historic linear mixed model in Equation (1). Second, the predicted random effects $\widehat{\boldsymbol{b}}_{it}$ and patient age are used as predictors in the historic logistic regression model of Equation (2) to provide at week t the predicted probability of developing GTN for this new patient, $\hat{\pi}_{it}$.

The above two steps are applied differently in the maximum likelihood and the Bayesian methods. In the maximum likelihood models, the random effects are predicted using the EBLUP, where the EBLUP of the random effects is defined as the mean of their conditional distribution given the repeated hCG measurements and age (Lindstrom & Bates, 1988). In the Bayesian models, random effects are predicted by taking the mean of the subject-specific random effects posterior distribution. For more illustration, see the scheme in Figure 3 and the R syntax in the supplementary material.

The above steps can be repeated to obtain an updated prediction $\hat{\pi}_{it}$ at any time a new measurement is recorded for this patient, by adjusting the $\boldsymbol{y}_{it}$ vector and the corresponding X_{it} and Z_{it} matrices. They can also be applied to obtain predictions for a patient with missing observations, by adjusting the design matrices to match the corresponding available measurements. An example of dynamic prediction calculations is given in Appendix A.

We explained so far how to obtain GTN-predicted probabilities; however, deriving their standard errors and prediction intervals is rather not straightforward since we need to account for the uncertainty of both the fixed effects and random effects estimates. To obtain the 95% prediction intervals for the $\hat{\pi}_{it}$ based on the likelihood approaches, we perform bootstrapping to fully account for all sampling uncertainty, as follows:

1. Bootstrap patients (so including their full biomarker trajectory) in the model development phase to obtain B sets (e.g., $B = 1000$) of the estimated model parameters for the historic model.
2. Obtain B updated GTN prediction probabilities $\hat{\pi}_{it}^{B}$ for the new patient (following the scheme in Figure 3).
3. Compute the updated prediction $\hat{\pi}_{it}$ and its corresponding 95% prediction interval by taking the percentiles of the probabilities $\hat{\pi}_{it}^{B}$ in Step 2.

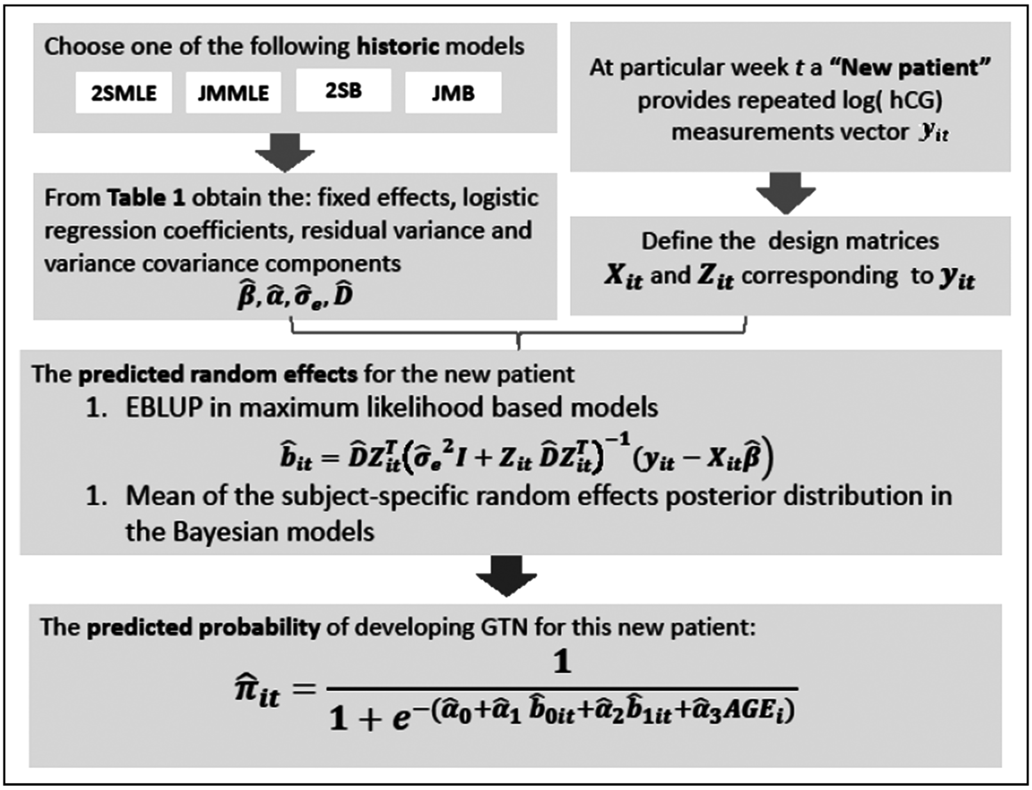

Choose one of the following **historic** models
2SMLE
JMMLE
2SB
JMB
At particular week *t* a **"New patient"** provides repeated log(hCG) measurements vector $\boldsymbol{y}_{it}$
From **Table 1** obtain the: fixed effects, logistic regression coefficients, residual variance and variance covariance components
$\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\sigma}}_{\boldsymbol{e}}, \hat{\boldsymbol{D}}$
Define the design matrices $\boldsymbol{X}_{it}$ and $\boldsymbol{Z}_{it}$ corresponding to $\boldsymbol{y}_{it}$
The **predicted random effects** for the new patient
1. EBLUP in maximum likelihood based models
$\hat{\boldsymbol{b}}_{it} = \hat{\boldsymbol{D}}\boldsymbol{Z}_{it}^{T}(\hat{\boldsymbol{\sigma}}_{\boldsymbol{e}}{}^{2}\boldsymbol{I} + \boldsymbol{Z}_{it}\hat{\boldsymbol{D}}\boldsymbol{Z}_{it}^{T})^{-1}(\boldsymbol{y}_{it} - \boldsymbol{X}_{it}\hat{\boldsymbol{\beta}})$
1. Mean of the subject-specific random effects posterior distribution in the Bayesian models
The **predicted probability** of developing GTN for this new patient:
$\hat{\boldsymbol{\pi}}_{it} = \dfrac{1}{1 + e^{-(\hat{\alpha}_0 + \hat{\alpha}_1\hat{b}_{0it} + \hat{\alpha}_2\hat{b}_{1it} + \hat{\alpha}_3 AGE_i)}}$

FIGURE 3 The scheme for predicting GTN for a new patient using the four approaches

In the Bayesian approach, we apply the same steps but instead of bootstrapping, we directly use B samples from the posterior distribution of the model parameters, for more details, see the provided code in the supplementary material.

4 | RESULTS

4.1 | Model development

We applied the four approaches (2SMLE, 2SB, JMMLE, and JMB) using the GTN data set. Table 1 represents the parameter estimates of the four historic models with their corresponding 95% confidence and credible intervals. The coefficient for age, $\hat{\beta}_2$, in relation to hCG was estimated to be approximately zero by all approaches; hence, we dropped it from the longitudinal submodels. Comparing the estimates of the association parameters α_1 and α_2 in the four models, the joint models gave slightly higher than the two-stage models. The reason for that is attenuation, due to the ignored uncertainty in the estimated random effects, which was not carried forward to the binary submodel in the two-stage methods. The 95% confidence and credible intervals of α_1 and α_2 in the binary submodel show a significant association between the longitudinal hCG profile and the GTN risk in all approaches.

4.2 | The dynamic predictive performance

The dynamic predictive performance measures (AUC, MSEP, and MCER) using the four approaches are shown in Figure 4. Each measure was obtained using the four historic models to obtain the updated predictions per week, for example, the AUC at week 4 for the JMB reflects the predictive performance of the historic JMB to predict the probability of GTN, based on the available hCG measurements from week 2 until week 4 (i.e., based on three hCG measurements). The predictive performances for the four models were equivalently high, with a general trend of increase in predictive performance when the number of available hCG measurements increased.

In Figure 4, we notice that the rate of improvement in prediction accuracy starts to decrease after week 4, while the misclassification rate (MCER) starts with 20% misclassified patients, but continues to decrease almost constantly till week 5, where 10% of the patients are misclassified. This implies that adding measurements beyond week 4 or 5 will have only a small improvement on the accuracy of predictions. Table 2 represents the classification matrix for the four approaches using as an example the data available till week 4 and a cutoff of 50% for the risk of GTN. The MCER was around 0.125, which implies that 12.5% of the patients were classified in the wrong group.

4.3 | The internal validation of the dynamic prediction performance

The internal validation of the dynamic AUC(t)s and MSEP(t)s using bootstrap showed that the four models have very small to negligible optimism (<0.006 in the AUC and <0.01 in the MSEP), hence we have reported the nonadjusted measures, see Table B1 and Table B2 in Appendix B.

4.4 | Subject-specific dynamic prediction for new patients

We used the results in Table 1 to predict GTN for selected “new” patients, Patients A and B, with positive and negative GTN status, respectively, who were excluded from our data set in the base model fitting process to use them here.

We observe from the longitudinal trajectories of log(hCG) for these two patients in Figure 5, that Patient A showed persistently elevated hCG levels, which is indicative of a high risk of GTN. On the other hand, Patient B showed an increase in the log(hCG) levels at week 2 followed by gradual decline, and therefore we expect her eventually to be in the low-risk group.

The updated predictions of GTN for Patients A and B were obtained by applying the scheme in Figure 2 and by adjusting the matrices X and Z to match the available data in the corresponding weeks. More details on how to compute the updated predictions is illustrated for Patient B in Appendix A. Figure 6 presents the log(hCG) observations with the corresponding predicted GTN probabilities and 95% prediction intervals at each week, using the four models. The risk predictions for the two patients at week 2 start similarly in Panel a. In Panel b, we observe for Patient B an increase in the level of log(hCG) at week 3, which is reflected in a strong increase of her predicted GTN risk. After week 3, her log(hCG)

TABLE 1 Parameter estimates of the models predicting GTN using the weekly hCG measurements

		Two-Stage Model (2SMLE)			Bayesian Two-stage (2SB)			Maximum likelihood Joint Model (JMMLE)			Joint Bayesian Model (JMB)		
			95% CI[a]			95% CI[b]			95% CI[a]			95% CI[b]	
		Estimate	**Lower**	**Upper**	**Estimate**	**Lower**	**Upper**	**Estimate**	**Lower**	**Upper**	**Estimate**	**Lower**	**Upper**
Longitudinal submodel													
Fixed Intercept	$\boldsymbol{\beta}_0$	2.50	2.44	2.57	2.51	2.44	2.57	2.50	2.44	2.57	2.48	2.44	2.53
Fixed slope (week)	$\boldsymbol{\beta}_1$	−0.22	−0.24	−0.20	−0.21	−0.23	−0.2	−0.22	−0.24	−0.20	−0.22	−0.23	−0.19
SD for random intercept b_0	$\boldsymbol{d}_{11}$	0.59	0.54	0.65	0.59	0.54	0.65	0.59	0.53	0.64	0.58	0.53	0.64
SD for random slope b_1	$\boldsymbol{d}_{22}$	0.18	0.17	0.20	0.19	0.18	0.21	0.18	0.17	0.2	0.19	0.18	0.21
Random effects covariance	$\boldsymbol{d}_{12}$	−0.01	−0.02	0.00	−0.01	−0.03	0.00	−0.01	−0.02	0.00	−0.01	−0.03	0.00
Residual *SD*	$\boldsymbol{\sigma}_{\mathbf{e}}$	0.19	0.18	0.20	0.19	0.18	0.20	0.19	0.18	0.20	0.19	0.18	0.20
Binary submodel													
Intercept	$\boldsymbol{\alpha}_0$	−2.07	−3.59	−0.55	−1.75	−3.10	−0.54	−2.41	−4.21	−0.60	−2.43	−4.26	−0.71
Coefficient for b_0[c]	$\boldsymbol{\alpha}_1$	0.92	0.44	1.41	0.79	0.35	1.24	1.05	0.56	1.54	1.03	0.54	1.52
Coefficient for b_1[d]	$\boldsymbol{\alpha}_2$	3.82	2.97	4.68	3.67	3.02	4.43	4.58	3.38	5.78	4.37	3.45	5.61
Age	$\boldsymbol{\alpha}_3$	0.02	−0.02	0.067	0.02	−0.00	0.06	0.03	−0.03	0.08	0.03	−0.02	0.08

[a] Confidence interval.

[b] Credible interval.

[c] A change of 1 unit (unit = d_{11}) in $\boldsymbol{b}_0$ leads to $\boldsymbol{\alpha}_1$ units change in log(odds ratio).

[d] A change of 1 unit (unit = d_{22}) in $\boldsymbol{b}_1$ leads to $\boldsymbol{\alpha}_2$ units change in log(odds ratio).

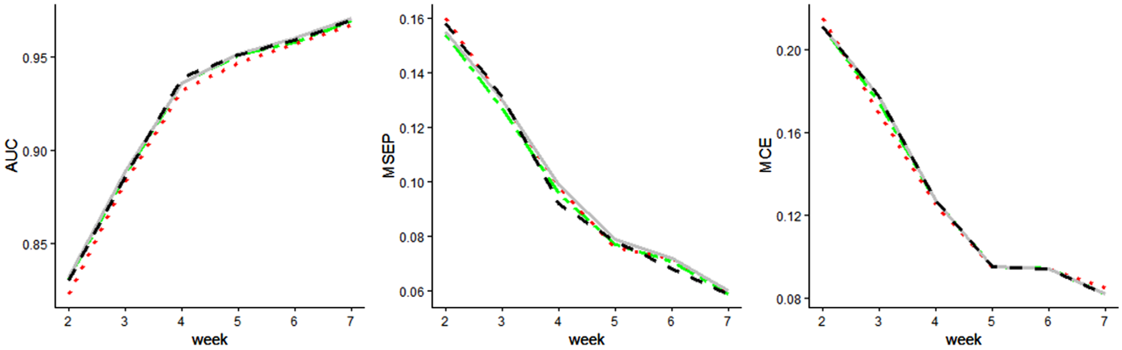


FIGURE 4 The area under the ROC curve (AUC) (left panel), the mean square error of prediction (middle panel) (MSEP), and the misclassification rate (MCER) (right panel) per week using the four different approaches

TABLE 2 The classification matrix of the patients based on the available hCG measurements till week 4 and using the four approaches

Predicted GTN Status									
	2SMLE		2SB		JMMLE		JMB		
Observed Status	No	Yes	No	Yes	No	Yes	No	Yes	Total
No GTN	250	11	249	12	249	12	249	12	261
GTN	37	87	37	87	37	87	37	87	124
Total	287	98	286	99	286	99	286	99	385
MCER	0.125		0.127		0.127		0.127		

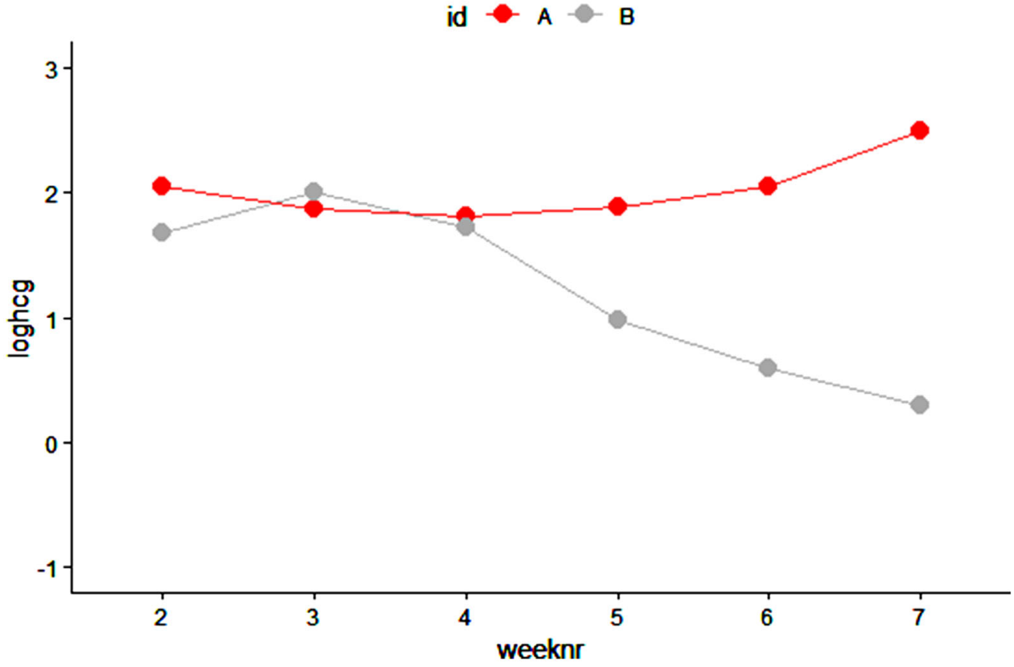


FIGURE 5 Observed longitudinal log(hCG) trajectories for Patients A and B

level starts to decrease, and her predicted GTN risk also decreases. Patient A, who has stable levels of log(hCG) followed by a slight increase, starts at week 2 with low GTN risk, which increases gradually as more information becomes available.

Comparing the updated predicted probabilities obtained from the four methods in Figure 6, the four methods have similar predicted values and prediction intervals in this example.

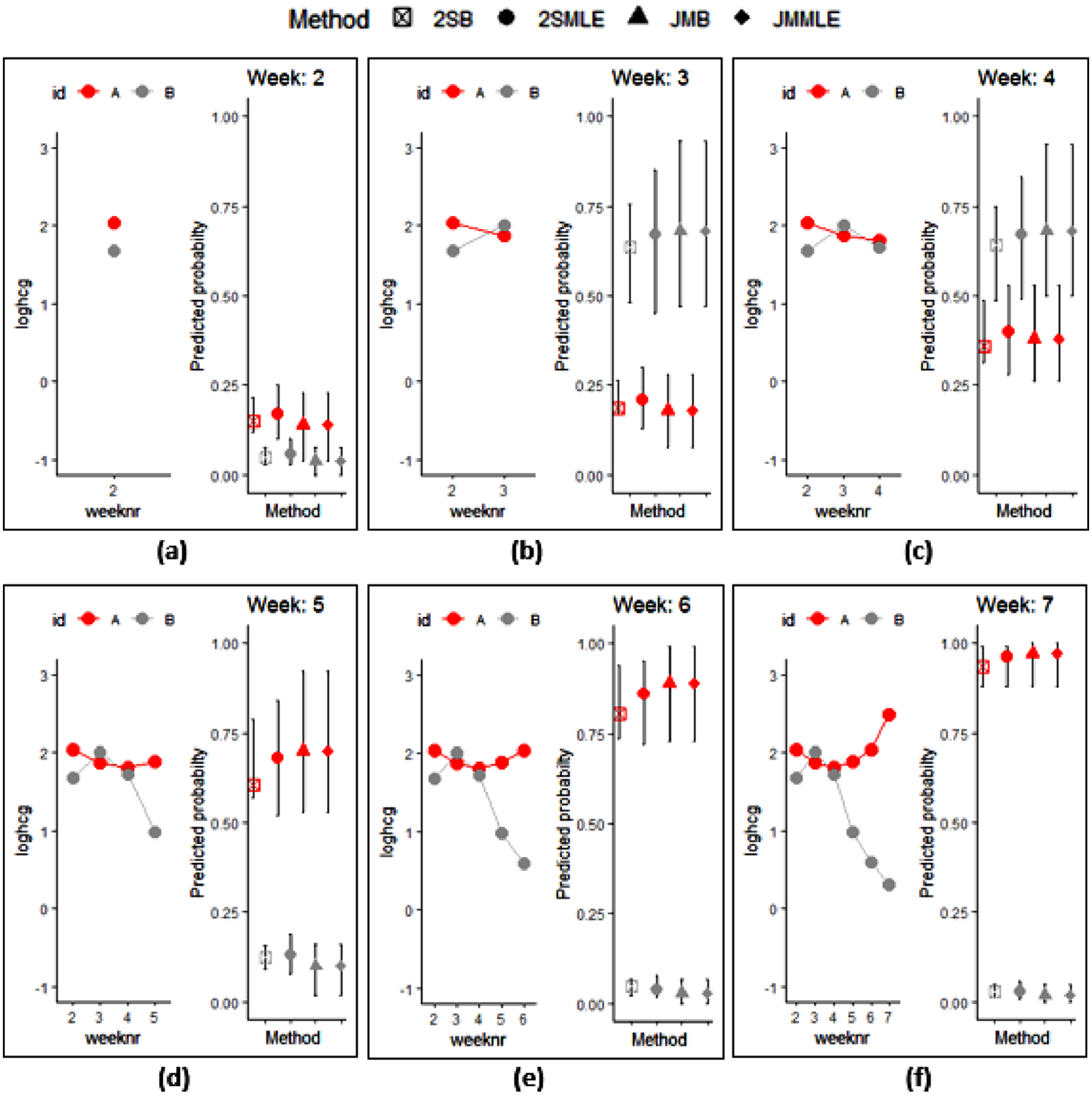


FIGURE 6 Dynamically updated predicted probabilities of developing post–molar GTN based on hCG for two selected patients using the four approaches. Left panels: observed log(hCG) measurements at each week. Right panels: the corresponding predicted probabilities and 95% prediction intervals. (a) week 2, (b) week 3, etc

5 | SIMULATION STUDY

Based on our results, despite the methodological differences between the two-stage methods and the joint models, they performed the same in predicting GTN. This could be due to the very small within- and between-patients variability of the repeatedly measured biomarker (hCG) in our example data set, which produced random effects from the longitudinal model that were apparently able to reflect all the needed information to properly predict GTN whether we performed the prediction in two separate stages or jointly in one model. Thus, the expected added value of the joint modeling technique was not detected by comparing the predictive performance measures. In response to this, a simulation study was needed to further investigate the obtained results.

In this section, we present the results of a simulation study where we simulated data that resemble the original GTN data set with the same number of patients, events and weeks, and the same percentage of missing values. The main purpose of our simulation is to investigate the sensitivity of the results that we obtained with the original GTN data when we increase the inter- and intrapatient variances, that is, the variance of the random error (which include measurements error and/or physiological variation over time) and the variance–covariance component of the subject-specific random effects.

TABLE 3 Description of the four scenarios used in simulating hCG profiles for GTN and non-GTN groups

			GTN patients = 140			Non-GTN patients = 299		
	Simulation Scenario	σ_ϵ^2	d_{11}	d_{22}	d_{12}	d_{11}	d_{22}	d_{12}
1	Simulated original	0.04	0.57	0.08	−0.01	0.63	0.15	−0.01
2	Increased between variability	0.04	1.14	0.16	−0.02	1.26	0.30	−0.04
3	Increased within variability	4.00	0.57	0.08	−0.01	0.63	0.15	−0.01
4	Increased within and between variability	4.00	1.14	0.16	−0.02	1.26	0.30	−0.04

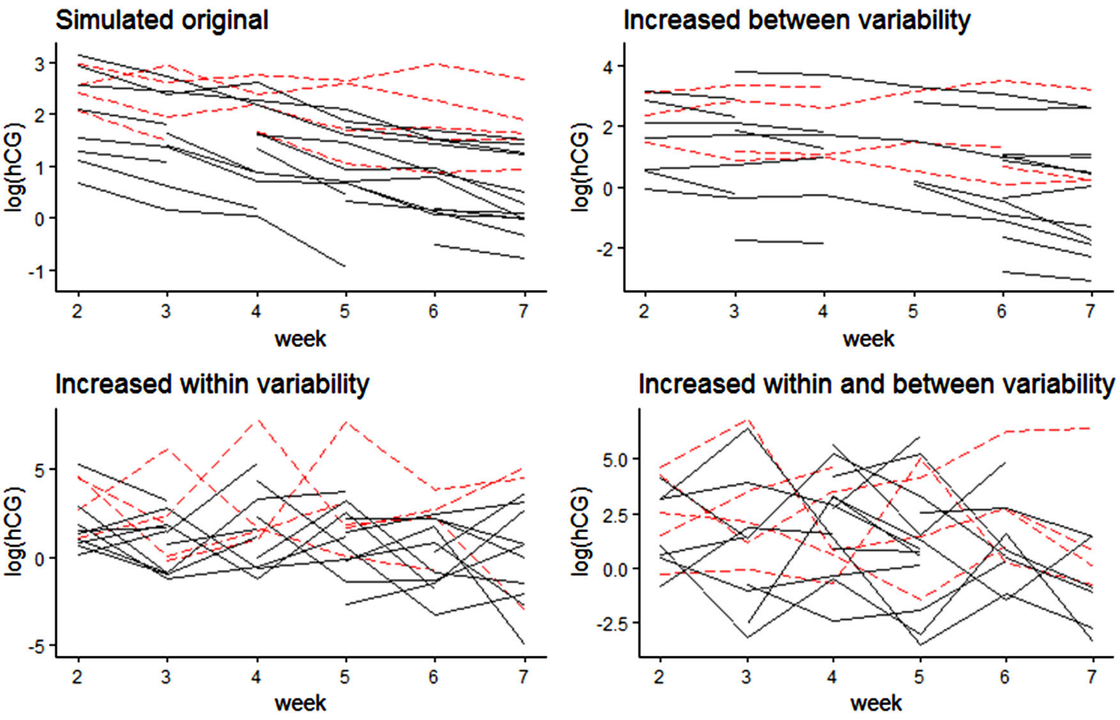


FIGURE 7 The log(hCG) profiles for the 50 randomly selected subjects from the four simulation scenarios (red = GTN patients, black = uneventful women)

We applied four different simulation scenarios to generate the data. In each scenario, the simulation was done by changing the inter- and intrasubject variances while retaining the other parameters as in the GTN data. Table 3 represents a description of the parameters associated with the four simulation scenarios. The four scenarios were chosen to represent different scenarios of bigger and smaller random error variances and bigger and smaller variation in the random effects.

The longitudinal log(hCG) profiles of random samples (50 patients) from the simulated data were generated using the four scenarios and are presented in Figure 7.

In order to keep it short, we obtained the predictive performance for the four models based on the measurements of all weeks, that is, from week 2 to week 7. Table 4 represents a summary of the predictive performances (the AUC and the MSEP) of the four modeling approaches using the four different simulation scenarios by increasing the inter- and intravariability of the log(hCG) measurements as described in Table 3. The predictive performance of the joint models was higher than that of the two-stage models in all the scenarios, with higher AUCs and lower MSEPs. However, this outperformance of the joint models was very noticeable when we increased the random error variance in Scenario 3 and even more noticeable with increasing both the random error variance and the variance–covariance component of the random effects in Scenario 4. On the other hand, increasing the variability resulted in deterioration in the predictive performance in the two-stage models, the worst predictive performance for the two-stage models was in Scenario 4 when we increased both the within- and between-patients variabilities; however, the joint models were more robust to changes in the within-patients variabilities than in changes to the between-patients variability.

TABLE 4 A summary of the predictive performances (the area under the ROC curves and the mean squared error of prediction) of the four modeling approaches using the four different simulations scenarios

		AUC				MSEP			
	Simulation Scenario	**2SB**	**JMB**	**2SMLE**	**JMMLE**	**2SB**	**JMB**	**2SMLE**	**JMMLE**
1	Simulated original	0.96	0.98	0.96	0.99	0.07	0.04	0.07	0.04
2	Increased between variability	0.80	0.83	0.80	0.85	0.15	0.14	0.15	0.13
3	Increased within variability	0.80	0.99	0.81	0.98	0.16	0.05	0.16	0.06
4	Increased within and between variability	0.72	0.90	0.72	0.91	0.19	0.13	0.19	0.12

6 | DISCUSSION AND CONCLUSION

In this article, we illustrated and compared different modeling approaches that allow longitudinal continuous biomarker profiles to be used as predictors to provide updated predictions for a binary outcome, namely, the two-stage and the joint modeling approach. We used two different estimation methods: maximum likelihood and Bayesian, and we illustrated their performance in predicting post–molar GTN.

The four modeling approaches (JMB, 2SMLE, 2SB, and JMMLE) have many advantages over the simpler methods that were discussed recently by Welten et al. (2018). For example, the two-stage method and the joint modeling approach can handle missing values in the repeated measurements, and they have no limitations to the minimum required number of measurements per patient or the irregularity of spacing between the repeated measurements. On the other hand, our approaches are equally flexible as the growth curve method that Welten et al. recommended (Welten et al., 2018), as both methods are capable of incorporating longitudinal data. Moreover, we showed that models based on random effects can be used for prediction, and need the same amount of calculations as their preferred growth curve method, even though they suggested that using random effects may not be practical.

An additional advantage of the four approaches is their flexibility to include additional covariates (baseline characteristics) to improve the accuracy of prediction. One could consider adding covariates in the longitudinal submodel if they are expected to be related to the evolution of the biomarker, while adding others to the logistic submodel if they are related to the probability of developing the outcome. Moreover, the Bayesian approaches allow the flexibility to vary parametric assumptions by the specification of prior distributions for the parameters, which gives us the opportunity to incorporate any existing knowledge into the model.

Despite the methodological differences between the two-stage methods and the joint models, the four methods (JMB, 2SMLE, 2SB, and JMMLE) performed the same in predicting GTN. They showed very good predictive accuracy when applied to the GTN data, meaning that patients with high risk of post–molar GTN could be accurately differentiated from those with low risk by using the available hCG measurements. The two-stage methods performed as good as the joint modeling approach for predicting post–molar GTN, even though they do not correct for uncertainty in the estimated random effects. The equivalent performance could be however due to the very small within- and between-patients variability of the repeatedly measured biomarker (hCG) in our GTN data, in addition to the good fit of the longitudinal model that yielded sufficiently precise estimates of the random effects that were able to reflect all the needed information to properly predict GTN. Therefore, the equivalence of the methods may not hold to other applications with higher variability.

The internal validation of the four models showed negligible optimism in the predictive performance, however, this may not be a guarantee for the good performance of the model if applied to another group of GTN patients. Therefore, additional validation is needed using an external data set.

Based on the results we obtained using both the original GTN data set and the simulated data sets, we recommend using the joint modeling approaches (JMB and JMMLE) in applications where the data are measured with high error or/and when the variability between the study subjects is high.

Comparing the computational cost of the four approaches, fitting the maximum likelihood models was much faster than the Bayesian ones, as expected due to the need for MCMC methods. The two-stage methods are easy to fit using standard regression models and available software. However, they need to be fitted in two steps, where you have to merge the results of the first stage with your original data to perform the second stage. On the other hand, joint modeling approaches are fitted in one stage and require less data management steps, but they require extra programming. This makes the two-stage methods a good and an easy alternative for researchers who do not have the needed technical skills; however, they should keep in mind that the two-stage methods might be not the optimal option in applications with high variability.

In this paper, we provided a toolbox of four models to obtain updated predictions in cases where longitudinal biomarkers are associated with future binary outcomes, which could enable better personalized decisions on interventions for patients. More generally, the toolbox is not only useful for predicting binary outcomes but also for categorical, count, and continuous outcomes based on biomarker data recorded over time, by replacing the binary submodel with the suitable corresponding regression model. Finally, to facilitate the applicability of our four approaches, we included in the supplementary material a detailed programming syntax. This is an important contribution in the absence of packages that fit this kind of models for binary outcomes.

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

ORCID

Rana Dandis https://orcid.org/0000-0003-4002-1386

REFERENCES

Aminimoghaddam, S., Yarandi, F., Nejadsalami, F., Taftachi, F., Noor Bakhsh, F., & Mahmoudzadeh, F. (2014). Human chorionic gonadotrophin as an indicator of persistent gestational trophoblastic neoplasia. *Medical Journal of the Islamic Republic of Iran*, *28*, 44.

Cavaliere, A., Ermito, S., Dinatale, A., & Pedata, R. (2009). Management of molar pregnancy. *Journal of Prenatal Medicine*, *3*(1), 15–17.

Emmanuel Lesaffre, A. B. L. (2013). *Bayesian biostatistics*. New York, NY: John Wiley & Sons.

Eysbouts, Y., Massuger, L., Ottevanger, P., IntHout, J., & Sweep, F. (2017). *Nomogram-based prediction of post-molar gestational trophoblastic neoplasia with serum human chorionic gonadotropin*. Nijmegen, the Netherlands: Radboud University Medical Centre. *Unpublished manuscript*.

Harrell, F. E., Jr., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, *15*(4), 361–387.

Horrocks, J., & van Den Heuvel, M. J. (2009). Prediction of pregnancy: A joint model for longitudinal and binary data. *Bayesian Analysis*, *4*(3), 523–538.

Ibrahim, J. G., Chu, H., & Chen, L. M. (2010). Basic concepts and methods for joint models of longitudinal and survival data. *Journal of Clinical Oncology*, *28*(16), 2796–2801.

Inc, S. I. (2015). *SAS/IML® 14.1 user's guide*. Cary, NC.

Khosravirad, A., Zayeri, F., Baghestani, A. R., Yoosefi, M., & Bakhtiyari, M. (2017). Predictive power of human chorionic gonadotropin in post-molar gestational trophoblastic neoplasia: A longitudinal ROC analysis. *International Journal of Cancer Management*, *10*(9):e9015. https://doi.org/10.5812/ijcm.9015

Kim, B. W., Cho, H., Kim, H., Nam, E. J., Kim, S. W., Kim, S., … Kim, J. H. (2012). Human chorionic gonadotrophin regression rate as a predictive factor of postmolar gestational trophoblastic neoplasm in high-risk hydatidiform mole: A case-control study. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *160*(1), 100–105.

Lindstrom, M. J., & Bates, D. M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, *83*(404), 1014–1022.

Mousavi, A. S., Karimi, S., Modarres Gilani, M., Akhavan, S., & Rezayof, E. (2014). Does Postevacuation beta: Human chorionic gonadotropin level predict the persistent gestational trophoblastic neoplasia? *Obstetrics & Gynecology*, *2014*, Article ID 494695.

Rizopoulos, D. (2012). *Joint models for longitudinal and time-to-event data: With applications in R*. New York: C. H. C. B. Series. 275 p.

Stevens, F. T., Katzorke, N., Tempfer, C., Kreimer, U., Bizjak, G. I., Fleisch, M. C., & Fehm, T. N. (2015). Gestational trophoblastic disorders: An update in 2015. *Geburtshilfe Frauenheilkd*, *75*(10), 1043–1050.

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Tsiatis, A. A., Degruttola, V., & Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS *Journal of the American Statistical Association*, *90*(429), 27–37.

Wang, C. Y., Wang, N., & Wang, S. (2000). Regression analysis when covariates are regression parameters of a random effects model for observed longitudinal measurements. *Biometrics*, *56*(2), 487–495.

Welten, M., de Kroon, M. L. A., Renders, C. M., Steyerberg, E. W., Raat, H., Twisk, J. W. R., & Heymans, M. W. (2018). Repeatedly measured predictors: A comparison of methods for prediction modeling. *Diagnostic and Prognostic Research*, *2*(1), 5.

Wulfsohn, M. S., & Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics*, *53*(1), 330–339.

SUPPORTING INFORMATION

Additional supporting information including source code to reproduce the results may be found online in the Supporting Information section at the end of the article.

How to cite this article: Dandis R, Teerenstra S, Massuger L, Sweep F, Eysbouts Y, IntHout J. A tutorial on dynamic risk prediction of a binary outcome based on a longitudinal biomarker. *Biometrical Journal*. 2020;62:398–413. https://doi.org/10.1002/bimj.201900044

APPENDIX A: SUBJECT-SPECIFIC DYNAMIC PREDICTION

We use Patient B as an example to illustrate the dynamic prediction calculations. We show how to calculate her GTN risk ($\hat{\pi}_{B,\ 3}$) based on the hCG observations for the weeks 2 and 3, using the JMMLE approach as an example. We follow the scheme in Figure 2. First, we use the results for $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{D}}$, and $\hat{\boldsymbol{\sigma}}_e$ from Table 1 combined with her log(hCG) measurements in $\boldsymbol{y}$ to predict her random effects $\hat{\boldsymbol{b}}_{B,3}$ at week 3 as follows:

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} 2.50 \\ -0.22 \end{bmatrix}, \quad \hat{D} = \begin{bmatrix} 0.35 & -0.01 \\ -0.01 & 0.03 \end{bmatrix}, \quad \hat{\sigma}_e^2 = (0.19)^2,$$

$$\boldsymbol{y} = \begin{bmatrix} 1.68 \\ 2.00 \end{bmatrix} \text{ and } X = Z = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix},$$

$$\hat{\boldsymbol{b}}_{B,3} = \hat{D}\, Z^T \left(\hat{\sigma}_e^2 I + Z\, \hat{D} Z^T\right)^{-1} \left(y - X\hat{\beta}\right),$$

$$\hat{\boldsymbol{b}}_{B,3} = \begin{bmatrix} \hat{b}_{0,B,3} \\ \hat{b}_{1,B,3} \end{bmatrix} = \begin{bmatrix} -0.41 \\ 0.13 \end{bmatrix}.$$

Next, $\hat{b}_{0,B,3}$ and $\hat{b}_{1,B,3}$, patient age ($AGE_B = 35$) and the binary submodel coefficients $\hat{\boldsymbol{\alpha}}$ from Table 1 are used to predict the risk of GTN as follows:

$$\hat{\boldsymbol{\alpha}} = \begin{bmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \end{bmatrix} = \begin{bmatrix} -2.41 \\ 1.78 \\ 24.10 \\ 0.03 \end{bmatrix},$$

$$\hat{\pi}_{B,3} = \frac{1}{1 + e^{-\left(\hat{\alpha}_0 + \hat{\alpha}_1 \hat{b}_{0,B,3} + \hat{\alpha}_2 \hat{b}_{1,B,3} + \hat{\alpha}_3 AGE_B\right)}},$$

$$\hat{\pi}_{B,3} = \frac{1}{1 + e^{-(-2.41 + (1.78 \times -0.41) + (24.10 \times 0.13) + (0.03 \times 35))}} = 0.74.$$

The predicted probability $\hat{\pi}_{B,3}$ is 0.74, meaning that after three weeks and by using two hCG measurements we predict at week 3, that Patient B is at 74% risk of developing GTN. The same calculations can be repeated to obtain GTN prediction whenever a new observation of hCG becomes available by adjusting the X and Z design matrices to the dimension of the hCG measurements vector y.

APPENDIX B

TABLE B1 The area under the ROC curve (AUC) of the updated prediction for each week using the four different approaches

	Bayesian Two-Stage (2SB)			Two-Stage Model (2SMLE)			Joint Bayesian Model (JMB)			Maximum likelihood Joint Model (JMMLE)		
Week	**AUC**	**95% CI**		**AUC**	**95% CI**		**AUC**	**95% CI**		**AUC**	**95% CI**	
2	0.831	0.822	0.834	0.823	0.814	0.829	0.832	0.828	0.834	0.830	0.822	0.831
3	0.887	0.878	0.890	0.883	0.871	0.888	0.889	0.885	0.890	0.886	0.877	0.888
4	0.936	0.929	0.938	0.932	0.926	0.934	0.936	0.934	0.938	0.939	0.935	0.941
5	0.951	0.946	0.952	0.947	0.942	0.948	0.952	0.950	0.952	0.951	0.947	0.952
6	0.958	0.954	0.960	0.957	0.955	0.959	0.960	0.957	0.960	0.959	0.955	0.960
7	0.970	0.965	0.971	0.967	0.964	0.969	0.971	0.969	0.971	0.970	0.966	0.971

TABLE B2 The mean squared error of the updated prediction for each week using the four different approaches

	Bayesian Two-Stage (2SB)			Two-Stage Model (2SMLE)			Joint Bayesian Model (JMB)			Maximum likelihood Joint Model (JMMLE)		
Week	**MSEP**	**95%CI**		**MSEP**	**95% CI**		**MSEP**	**95% CI**		**MSEP**	**95% CI**	
2	0.154	0.149	0.167	0.160	0.156	0.165	0.155	0.151	0.165	0.158	0.151	0.176
3	0.127	0.123	0.133	0.130	0.127	0.133	0.130	0.124	0.136	0.131	0.125	0.145
4	0.096	0.092	0.101	0.098	0.096	0.101	0.099	0.094	0.103	0.092	0.087	0.101
5	0.077	0.074	0.080	0.076	0.075	0.078	0.079	0.076	0.082	0.078	0.075	0.083
6	0.071	0.069	0.074	0.072	0.070	0.074	0.072	0.070	0.075	0.068	0.065	0.075
7	0.059	0.057	0.062	0.060	0.058	0.060	0.060	0.058	0.063	0.059	0.058	0.063