computation

MDPI

Article

Exploring Flexible Penalization of Bayesian Survival Analysis Using Beta Process Prior for Baseline Hazard

Kazeem A. Dauda [1,*], Ebenezer J. Adeniyi [2], Rasheed K. Lamidi [2] and Olalekan T. Wahab [2]

[1] Department of Mathematics, University of Bergen, 5007 Bergen, Norway
[2] Department of Mathematics and Statistics, Kwara State University, Malete, P.M.B. 1530, Ilorin 23431, Kwara State, Nigeria; ebenezerade247@gmail.com (E.J.A.); rasheedlamidi44@gmail.com (R.K.L.); taofeek.wahab@kwasu.edu.ng (O.T.W.)
* Correspondence: kazeem.dauda@uib.no or qdauda70@gmail.com

Abstract: High-dimensional data have attracted considerable interest from researchers, especially in the area of variable selection. However, when dealing with time-to-event data in survival analysis, where censoring is a key consideration, progress in addressing this complex problem has remained somewhat limited. Moreover, in microarray research, it is common to identify groupings of genes involved in the same biological pathways. These gene groupings frequently collaborate and operate as a unified entity. Therefore, this study is motivated to adopt the idea of a penalized semi-parametric Bayesian Cox (PSBC) model through elastic-net and group lasso penalty functions (PSBC-EN and PSBC-GL) to incorporate the grouping structure of the covariates (genes) and optimally perform variable selection. The proposed methods assign a beta process prior to the cumulative baseline hazard function (PSBC-EN-B and PSBC-GL-B), instead of the gamma process prior used in existing methods (PSBC-EN-G and PSBC-GL-G). Three real-life datasets and simulation scenarios were considered to compare and validate the efficiency of the modified methods with existing techniques, using Bayesian information criteria (BIC). The results of the simulated studies provided empirical evidence that the proposed methods performed better than the existing methods across a wide range of data scenarios. Similarly, the results of the real-life study showed that the proposed methods revealed a substantial improvement over the existing techniques in terms of feature selection and grouping behavior.

Keywords: penalization function; semi-parametric model; shrinkage parameter; joint posterior function; MCMC algorithm

Academic Editor: Dick De Ridder

Received: 20 October 2024
Revised: 23 December 2024
Accepted: 10 January 2025
Published: 21 January 2025

Citation: Dauda, K.A.; Adeniyi, E.J.; Lamidi, R.K.; Wahab, O.T. Exploring Flexible Penalization of Bayesian Survival Analysis Using Beta Process Prior for Baseline Hazard. *Computation* **2025**, *13*, 21. https://doi.org/10.3390/computation13020021

1. Introduction

Survival analysis is dedicated to investigating the time until a specific event, commonly termed as a "failure" [1]. This scenario finds widespread application across diverse scientific domains, including medicine [2,3], biology [3,4], engineering [5], and economics [6,7]. Researchers employ survival analysis to scrutinize the time leading up to events such as mortality or equipment malfunction [8]. For instance, researchers may investigate the survival times of cancer patients from the time of diagnosis, where the event of interest could be the occurrence of relapse or, unfortunately, the death of the patient [3,8–10].

Recently, there has been increasing interest in modeling survival data using deep learning methods in medical research [11–13]. However, these deep neural network-based survival models provide only point estimates of the hazard rates and thus cannot properly convey uncertainty in the estimations. Hence, to properly consider the uncertainties in deep neural network-based survival models, several researchers have proposed a Bayesian

deep learning approach [5,14]. As an illustration, Feng, and Zhao [14] proposed a Bayesian hierarchical deep neural network model that can provide not only a point estimate of survival probability but also quantification of the corresponding uncertainty. In addition, ref. [5] handled censored data and captured the non-linear relationship between covariates and target variable lifetime.

At the same time, Bayesian approaches to variable selection have also become popular, not least since the relevance of a covariate can be assessed simply by computing the posterior probability that it is included in a model [15,16]. This recasts variable selection as a model selection problem, with every possible model assigned an individual posterior probability. One of the most popular forms of such model selection priors is the spike-and-slab prior [2,17]. Among of the advantages of the method is the ability to adjust concentration points for flexible prior distributions, resulting in the nonlinear shrinkage of regression coefficients and smoothed estimates. In addition, it enables fully Bayesian inference through Markov Chain Monte Carlo, and aims to create a unified framework for variable selection, producing sparse coefficient structures in both low- and high-dimensional contexts [18–21].

Interestingly, among of the several traditional modeling methods for survival analysis well documented in the literature, the Cox proportional hazard (Cox-PH) model stands out as a prominent semi-parametric technique [22,23]. And, recognized for its practical applicability, it facilitates a flexible modeling approach with minimal assumptions by maximizing the partial likelihood and bypassing the need to model the baseline hazard function. In contrast, the Bayesian paradigm mandates an explicit parametrization of the baseline hazard function. However, extensive research has demonstrated the adoption of the gamma process [24–26], a prevalent non-parametric process prior, in Bayesian proportional hazard models to characterize the cumulative baseline hazard.

Similarly, Ibrahim et al. [25] proposed a semi-parametric approach, combining a non-parametric prior for the baseline hazard rate (using a discrete gamma process prior for the baseline hazard) with a fully parametric prior for the regression coefficients. Notably, the method demonstrated effective performance with low-dimensional data [25,27–29]. However, the methodology has yet to be applied to high-dimensional data, and its suitability for grouped data has not been explored.

Subsequently, while there have been studies on variable selection methods, ref. [30–34] pioneered a variable selection approach incorporating the widely recognized lasso penalty within the Cox proportional hazard (Cox-PH) model framework, where the likelihood function is grounded, and the cumulative baseline hazard function is modeled using a gamma process. The study introduced a prior on the tuning parameter for shrinkage, offering adaptive control over the model's sparsity. Similarly, Lee et al. [35] applied a data augmentation approach to handle censored survival times and enhance prior–posterior conjugacy. Also, for identifying relevant grouped covariates, a shrinkage prior distribution for regression coefficients, emulating the effect of a group lasso penalty, was assigned. And lastly, to address the challenge of not shrinking coefficient estimates to exact zeros in a Bayesian penalized regression approach, a two-stage thresholding method utilizing the scaled neighborhood criterion and Bayesian information criterion was employed. However, the method was found to outperform the competing methods on both simulated and real-life data, in terms of variable selection accuracy and predictive power.

Hence, in this paper, considering the typical distribution patterns of survival analysis, which often manifest as either gamma or beta distributions, the researchers have a reasonable basis to explore the incorporation of a beta process prior. Many researchers [34,36] have considered using a gamma process prior in Bayesian non-parametric settings due to its flexibility to adapt to various models; however, its relative complexity presents significant challenges for principled inference [37]. This is in contrast to the beta process

prior, which is mathematically tractable and serves as a fully Bayesian conjugate prior, enabling analytical posterior calculation and straightforward inference [38,39]. Thus, there is a need to implement the beta perspective within the framework, representing a modification of the methodologies proposed by [34,36]. Hence, this study implements and adopts the beta prior for the baseline hazard for effective prediction and variable selection in high-dimensional survival data. A similar version of the beta prior was introduced in the master's thesis of one of the authors of this study in 2023 [40]. In this phase, we present the mathematical formulation of the beta prior and the joint posterior for both elastic and group lasso, as well as real-life applications, respectively.

This paper is structured as follows: In Section 2.1, we present a comprehensive overview of the penalized semi-parametric Bayesian Cox model, incorporating grouping and shrinkage priors. Formulation of the modified joint posteriors is discussed in Section 2.2, while Section 2.3 provides a thorough computational algorithm for obtaining both the existing and proposed penalized semi-parametric Bayesian Cox model (PSBC). To assess the performance of our proposed methods, Section 3.1 present the simulation of this study. In Section 3.2, we apply the proposed models (PSBC-EN-B and PSBC-GL-B) alongside the existing models to analyze three distinct microarray gene expression datasets. Concluding this study, Section 4 presents our closing remarks on this study.

2. Methodology

In this section, we briefly review the general concept of the penalized semi-parametric Bayesian Cox model, examining existing methods and introducing new methods that focus on the prior of the cumulative baseline hazard function, joint posterior distributions, and the full Bayesian technique for selecting tuning parameters. Moreover, we discuss the formulation of the sequential BIC as a thresholding method for both the existing and proposed approaches.

2.1. Penalized Semi-parametric Bayesian Cox Model with Grouping and Shrinkage Priors

Suppose that a dataset consists of n subjects, and for the individual i^{th} subject we record the actual survival time T_i, covariates $X_i = (X_{i1}, \ldots, X_{ip})^{'}$, and the event indicator $\delta_i \in \{0, 1\}$. In right censoring, $T_i = min(T_i^c, C_i)$, where T_i^c *and* C_i are the survival and censoring times, respectively, $(i = 1, \ldots, n)$, and non-negative random variables, and the event indicator $\sigma_i = I(T_i^c \leq C_i)$, where $I(.)$ is an indicator function. The data structure for this scenario is properly illustrated in Table 1.

Table 1. Survival times, event indicators, and covariates for all subjects.

Subject	Survival Time T	Event Indicator δ	Covariates X_i
1	t_1	δ_1	$(x_{11}, x_{12}, \ldots, x_{1p})$
2	t_2	δ_2	$(x_{21}, x_{22}, \ldots, x_{2p})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	t_n	δ_n	$(x_{n1}, x_{n2}, \ldots, x_{np})$

Assume that the subjects are independent from each other and that $T_i^c \perp C_i$, given X_i. Thus, the conditional hazard function given by X_i $(h(t/X_i))$ quantifies the instantaneous failure rate at a given time t with X_i. Hence, the model that links the conditional hazard function to X_i is denoted by Equation (1), commonly called the Cox model [22], a well-known semi-parametric model in survival analysis.

$$h(t|x) = h_0(t)exp(x'\beta), \tag{1}$$

where $\beta = (\beta_1, \ldots, \beta_p)'$ is a column vector of p regression parameters and $h_0(t)$ is unspecified arbitrary baseline hazard function. Under the Cox model (1), the joint survival probability of n subjects given the matrix of covariates X is given by

$$P(T > t|\beta, X, H_0) = exp\left\{ -\sum_{i=1}^{n} exp(x_i'\beta) H_0(t_i) \right\}, \tag{2}$$

where $H_0(t_i)$ is the cumulative baseline hazard function. A gamma process prior is assigned to the cumulative baseline hazard function $H_0(t)$

$$H_0(t) \sim \mathcal{GP}(c_0 H^*, c_0). \tag{3}$$

$H^*(t)$ is an increasing function with $H^*(0) = 0$ and c_0 is a positive constant. Let h_j denote the increment in the cumulative baseline hazard in the interval I_j, as follows

$$h_j = H_0(s_j) - H_0(s_{j-1}), \ j = 1, 2, \ldots, J. \tag{4}$$

The gamma process prior in (4) implies that h_j follows an independent gamma distribution, that is

$$h_j = \mathcal{G}(\alpha_{0j} - \alpha_{0j-1}, c_0), \tag{5}$$

where $\alpha_{0j} = c_0 H^*(s_j)$. Therefore, the conditional probability of the i^{th} subject failing in the interval I_j is given by

$$\begin{aligned} P(T_i \in I_j|\mathbf{h}) &= exp(-exp(x_i'\beta) H(t_{j-1})) - exp(exp(x_i'\beta) H(t_i)) \\ &= exp(-exp(x_i'\beta) \sum_{m=1}^{j-1} h_m)(1 - exp(-h_j exp(x_i'\beta))), \end{aligned} \tag{6}$$

where $h = (h_1, h_2, \ldots, h_j)'$. This leads to our grouped data likelihood function

$$L(\beta, \mathbf{h}|D) \propto \prod_{j=1}^{J} G_j, \tag{7}$$

where $G_j = exp\{-h_j \sum_{k \in R_j - D_j} exp(x_i'\beta)\} \prod_{l \in D_j} \left[1 - exp\left\{ -h_j exp(x_i'\beta) \right\}\right]$.

The regression coefficients $\beta = (\beta_1, \ldots, \beta_p)$ play a major role in selecting the covariates in the model (7). So we introduce a prior that will result in shrinkage in our model. The variable selection is carried out through BIC thresholding. Note that the $|\beta_j|$ in the lasso penalty is proportional to the (minus) log-density of the Laplace distribution. The Laplace prior (8) for the regression coefficients is given by

$$\pi(\beta) = \prod_{j=1}^{p} \frac{\lambda}{2} exp(-\lambda|\beta_j|). \tag{8}$$

Meanwhile, we use the conditional Laplace prior in (9), instead of (8), in order to guarantee unimodality.

$$\pi(\beta|\sigma^2) = \frac{\lambda}{2\sqrt{\sigma^2}} exp\left(-\frac{\lambda|\beta|}{\sqrt{\sigma^2}} \right). \tag{9}$$

A noninformative marginal prior $\pi(\sigma^2) = \frac{1}{\sigma^2}$ is assigned on σ^2. The hierarchical representation of the Bayesian lasso with prior (9) can thus be obtained by utilizing the representation

of the Laplace distribution as a scale mixture of normals with an exponential mixing density, as described in (10).

$$\frac{a}{2}e^{-a|z|}\int_0^\infty \frac{1}{\sqrt{2\pi s}}e^{-\frac{z^2}{2s}}\cdot\frac{a^2}{2}e^{-\frac{a^2 s}{2}}ds. \tag{10}$$

The comprehensive details and theoretical foundation of the existing methods are discussed in the studies by [34,36].

2.2. Formulation of Proposed Methods: PSBC-EN-B and PSBC-GL-B Joint Posteriors

This section presents the formulation of the joint posterior of the PSBC-EN-B and PSBC-GL-B models. These models are constructed upon the grouped data likelihood (11), wherein the cumulative baseline hazard function in the Cox model is a priori modeled by a beta process. Moreover, the joint posterior of each model is based on the full Bayesian approach for selecting the tuning parameters.

However, these posteriors aim to improve the accuracy and efficiency of parameter estimation in high-dimensional statistical modeling, especially in scenarios characterized by complex data structures and a plethora of predictors. Additionally, through the incorporation of grouping and shrinkage priors, this modification seeks to strike a balance between variable selection, coefficient estimation, and computational tractability.

2.2.1. Development of PSBC Elastic Net (PSBC-EN-B) Joint Posterior

Assigning a beta process prior to the cummulative baseline hazard function $H_0(t)$, the likelihood can be written as

$$L(\beta, h|D) \propto \prod_{j=1}^{J} B_j, \tag{11}$$

where $B_j = \left((1-h_j)^{\Sigma exp(x_i'\beta)}\right)\prod_{l\in D_j}[1-(1-h_j)^{exp(x_i'\beta)}]$.

Here, $h = (h_1, h_2, \ldots, h_j)'$. To complete the discretized beta process model, we specify independent beta priors for the h_js. Specifically, we take

$$h_j \sim \mathcal{B}(c_{0j}\alpha_{0j}, c_{0j}(1-\alpha_{oj})) \tag{12}$$

as independent for $j = 1, 2, \ldots, J$.

Hence, the final joint posterior distribution of our `PSBC-EN-B` model can be written as

$$\begin{aligned}\pi(\beta, h, \sigma^2, \tau, \lambda_1^2, \lambda_2|D) &\propto L(D|\beta, h)\times\pi(h)\times\pi(\beta|\sigma^2,\tau,\lambda_2)\times\pi(\tau|\lambda_1^2,\lambda_2)\times\pi(\sigma^2)\\ &\quad\times\pi(\lambda_1^2)\times\pi(\lambda_2)\\ &\propto \prod_{j=1}^{J}\left((1-h_j)^{\sum_{k\in\mathcal{R}_j-\mathcal{D}_j}exp(x_i'\beta)}\right)\prod_{l\in D_j}[1-(1-h_j)^{exp(x_i'\beta)}]\\ &\quad\times\prod_{j=1}^{J}\left[h_j^{\alpha_{0j}\alpha_{0,j}-1}(1-h_j)^{c_{0j}(1-\alpha_{0j})-1}\right]|\sigma^2 D_\tau^*|^{-\frac{1}{2}}exp\left\{-\frac{1}{2\sigma^2}\beta' D_\tau^{*-1}\beta\right\}\\ &\quad\times\frac{1}{\sigma^2}(\lambda_1^2)^{r_1-1}exp(-\delta_1\lambda_1^2)\lambda_2^{r_2-1}exp(-\delta_2\lambda_2)\\ &\quad\times\prod_{j=1}^{p}\frac{1}{\sqrt{1+\lambda_2}\tau_j^2}exp\left(-\frac{\lambda_1^2\tau_j^2}{2}\right)C_1(\lambda_1^2,\lambda_2).\end{aligned} \tag{13}$$

2.2.2. Development PSBC Group Lasso (PSBC-GL-B) Joint Posterior

For the joint posterior distribution for the PSBC Group Lasso (PSBC-GL-B) model, given the cumulative baseline hazard function $H_0(t)$, the beta process prior can be expressed as follows:

$$\begin{aligned}\pi(\beta,h,\sigma^2,\tau_G,\lambda_2|D) &\propto L(D|\beta,h)\times\pi(h)\times\pi(\beta|\sigma^2,\tau_G)\times\pi(\tau_G|\lambda_2)\times\pi(\sigma^2)\times\pi(\lambda_2)\\ &\propto\prod_{j=1}^{J}\left((1-h_j)^{\sum_{k\in\mathcal{R}_j-\mathcal{D}_j}exp(x_i'\beta)}\right)\prod_{l\in D_j}[1-(1-h_j)^{exp(x_i'\beta)}]\\ &\times\prod_{j=1}^{J}\left[h_j^{\alpha_{0j}\alpha_{0,j}-1}(1-h_j)^{c_{0j}(1-\alpha_{0j})-1}\right]\prod_{k=1}^{K}(\sigma^2\tau_k^2I_{mk})^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}||\beta_k||^2_{G_k}\right\} \quad (14)\\ &\times\prod_{k=1}^{K}(\lambda^2)^{(m_k+1)/2}\tau_k^{(m_k+1)/2}\exp\left(-\frac{\lambda^2\tau_k^2}{2}\right)\\ &\times\frac{1}{\sigma^2}(\lambda^2)^{r-1}\exp(-\delta\lambda^2).\end{aligned}$$

2.3. Computational Scheme

This study adopts the Markov Chain Monte Carlo (MCMC) algorithm to fit our proposed models (PSBC-EN-B, Equation (13), and PSBC-GL-B, Equation (14)), described in the previous section, since the parameters in Equations (13) and (14) cannot be estimated analytically and thus require computational approaches. The computational approach in this study is similar to the method adopted by [34,41], and the MCMC simulation is used for variable selection, parameter estimation, and prediction in our proposed models (PSBC-EN-B, Equation (13), and PSBC-GL-B, Equation (14)). We refer the reader to [34] for full details of this computational scheme. However, the steps in Algorithms 1 and 2 depict the step-by-step MCMC algorithms adopted in fitting both the existing and proposed penalized semi-parametric Bayesian Cox model (PSBC).

Algorithm 1: MCMC algorithm step-by-step procedure for elastic-net

Data: Initial values: $\beta^{(0)}, h^{(0)}, \tau^{(0)}, \sigma^{2(0)}, \lambda_1^{2(0)}, and \lambda_2^{(0)}$. The index g denote the g^{th} iteration, and we start by setting $g = 1$

1 Set $g = 1$;
2 **while** $g < M$ **do**
3 Update β using Metropolis-Hastings algorithm with adaptive jumping rule;
4 Sample h_j;
5 from a gamma distribution (see Equation (5)) or beta distribution (using Equation (12)) Sample $1/\tau_j^2$;
6 Sample σ^2 from its full conditional distribution (inverse-gamma);
7 Update λ_1^2 using Metropolis-Hastings algorithm and the full conditional for λ_2;
8 Update λ_2 using Metropolis-Hastings algorithm and the full conditional for λ_2;
9 Increment $g = g + 1$;
10 **end**
11 Calculate the posterior mean and median of β to obtain estimators;

After obtaining the posterior estimates for each method, the sequential BIC were computed using Equation (15) below. The Bayesian information criterion (BIC) was then used as a thresholding method for variable selection in both the proposed and existing models.

Algorithm 2: MCMC algorithm step-by-step procedure for group lasso

Input : Initial values, $\beta^{(0)}, h^{(0)}, \tau^{(0)}, \sigma^{2(0)}$, and $\lambda^{2(0)}$. The index g denote the g^{th} iteration, and we start by setting $g = 1$

Output: Posterior samples of parameters

Set $g = 1$;
while $g \leq M$ *(M is the number of posterior samples desired)* **do**
    Update β using Metropolis-Hastings algorithm with adaptive jumping rule;
    Sample h_j's from a gamma distribution (see Equation (5)) or beta distribution (using Equation (12));
    Sample $1/\tau_k^2$ from inverse-Gaussian distribution;
    Sample σ^2 from inverse-gamma distribution;
    Sample λ^2 from gamma distribution;
    $g = g + 1$;
end
Calculate posterior mean and median of β to obtain estimators;

2.4. Bayesian Information Criterion (BIC)

The Bayesian information criterion (BIC) was adopted in this study to choose the variables for the models. After a model's absolute value estimations of β_j are sorted in descending order, the BIC values are calculated step by step by adding significant covariates in Equation (15).

$$BIC_j = -2\{l_j(\hat{\beta}_{(1:j)}) - l_0\} + j\log(n), \qquad j = 1, \ldots, p, \tag{15}$$

where $l_j(\hat{\beta}_{(1:j)})$ is the maximized log likelihood under a model $\mathcal{M}_j$ that includes covariates corresponding to the largest $j|\beta_j|$'s given by $(\hat{\beta}_{(1:j)})$, and $l_0(0)$ is the log likelihood under the null model. The lowest point on the BIC curve refers to the best choice of covariates.

3. Results

3.1. Simulation Study

In this section, we discuss simulation studies that were carried out to compare the performance of the methods. A simulation scenario, where a group of relevant covariates are correlated with others while the rest of the covariates are independent of each other, was designed.

Simulation Procedure: A $\rho = 0.5$ is used to generate the dataset. The 10 significant variables, whose coefficients are set equal to 4, are arranged near one another. This study assumes a pairwise correlation, with $\rho = 0.5$ only for the 10 important variables. The other $p - 10$ variables are assumed to be independent of each other. A dataset is generated using $\rho = 0.5$. The 10 significant variables, with coefficients set to 4, are positioned close to each other. This study assumes pairwise correlations of $\rho = 0.5$ for these 10 important variables only, while the remaining $p - 10$ variables are considered independent of each other.

Table 2 reports the Bayesian information criteria (BIC) and the average ranks of the BIC values across the PSBC models. The results reveal that `PSBC-GL-B` with a beta prior gives very impressive results by outperforming the competing models in terms of variable selection accuracy. Additionally, in almost all the cases the PSBC model (`PSBC-GL-B`) consistently gives the lowest BIC values (Table 2) compared to all other models.

Table 2. Simulation results based on 100 replications, with $\rho_{x_i,x_j} = 0.5$ and the average rank (AR) of the BIC.

Prior	Method	20	80	100	200	300	400	500	600	700	800	1000	AR
Gamma	PSBC-GL-G	2nd	2nd	2nd	1st	1st	2nd	2nd	2nd	2nd	2nd	2nd	
		−44.3	−230.1	−362.2	−889.3	−1175.4	−1564.9	−1941.1	−2181.7	−2779.4	−3154.9	−3817.5	1.81
	PSBC-EN-G	4th	3rd	3rd	4th	4th	3rd	3rd	4th	4th	3rd	4th	
		−28.5	−224.7	−264.9	−598.5	−939.7	−1482.8	−1614.6	−2047.3	−2163.8	−2585.6	−3032.5	3.54
Beta	PSBC-GL-B	1st	1st	1st	2nd	2nd	1st	1st	2nd	1st	1st	1st	
		−44.3	−269.3	−365.5	−862.9	−1120.7	−1567.3	−1971.2	−2186.8	−2862.1	−3171.8	−3842.5	**1.18**
	PSBC-EN-B	3rd	4th	4th	3rd	3rd	4th	4th	3rd	3rd	4th	3rd	
		−36.9	−197.4	−252.4	−611.7	−971.7	−1271.4	−1507.4	−2118.7	−2387.9	−2428.3	−3153.4	3.45

Overall, judging by the AR of the method based on the BIC values for each method across the selected number of predictors, we discovered that `PSBC-GL-B` with a beta prior was ranked first, and `PSBC-GL-G` with *a gamma prior* ranked second, followed by `PSBC-EN-B` *beta prior* and `PSBC-EN-G` *gamma prior*. Table 3 depicts the number of predictors selected by the `PSBC` models.

Table 3. The number of genes selected by PSBC models, using the BIC.

Cases	PSBC-GL-G	PSBC-GL-B	PSBC-EN-G	PSBC-EN-B
20	9	10	10	10
80	8	9	7	12
100	12	13	6	10
200	6	6	10	10
300	3	3	8	2
400	3	1	7	4
500	9	1	5	8
600	12	5	13	2
700	10	9	4	2
800	10	4	9	13
1000	21	15	15	16
Average Selection	9.3	**6.9**	8.5	8.1

Table 3 reveals the number of variables or predictors selected by the methods for each number of predictors in the data simulated. The last row of the table indicates the average number of predictors selected by the methods. The result indicate that modified `PSBC-GL-B` performed reasonably well compared to the competing methods.

To further demonstrate the performance of the proposed models, we considered using four operating characteristics, as has been applied in similar previous studies ([34–36]). The operating characteristics considered include the true positive rate (TPR or sensitivity), true negative rate (TNR or specificity), time-dependent ROC curve (AUC) [42,43], and the concordance index (c-index) [44].

The results presented in Table 4 summarize the overall averages and standard deviations of the four metrics for both the proposed and existing methods. From Table 4, it is evident that the PSBC-EN-B model outperforms the existing methods in terms of c-index accuracy. However, when considering time-dependent ROC curve (AUC) accuracy, both the proposed and existing methods appear to perform competitively. This observation suggests that as survival time increases, the hazard ratio distribution may exhibit properties of both gamma and beta distributions, influenced by patient demographic history (such as age, gender, etc.). Moreover, with regard to TPR accuracy, we observed equal performance across all models, while in terms of TNR, the existing methods were found to be relatively sensitive. For interested readers, we provide a visual presentation of the average AUC over time across various numbers of variables in the Supplementary Material (Figures S1–S4).

Table 4. Results of the average performance metrics of the simulation scenarios. The metrics are reported as percentages (%), with standard deviations provided in parentheses.

No. of Variables (P)	Method	TPR % (SD)	TNR % (SD)	AUC	C-INDEX
20	PSBC-GL-G	67.00 (0.295)	40.40 (0.301)	0.770	0.720
	PSBC-GL-B	60.80 (0.296)	**44.60** (0.302)	0.664	0.612
	PSBC-EN-G	**79.90** (0.307)	30.70 (0.335)	**0.997**	0.970
	PSBC-EN-B	**79.90** (0.307)	30.70 (0.335)	0.996	**0.971**
200	PSBC-GL-G	60.90 (0.284)	43.20 (0.307)	0.678	0.638
	PSBC-GL-B	58.40 (0.303)	**44.70** (0.291)	0.638	0.598
	PSBC-EN-G	80.40 (0.307)	30.70 (0.334)	**1.000**	0.982
	PSBC-EN-B	**80.60** (0.309)	30.50 (0.336)	**1.000**	**0.989**
500	PSBC-GL-G	54.20 (0.323)	**45.90** (0.284)	0.583	0.575
	PSBC-GL-B	61.30 (0.277)	44.00 (0.311)	0.674	0.630
	PSBC-EN-G	**80.30** (0.307)	30.90 (0.335)	0.999	0.980
	PSBC-EN-B	80.10 (0.308)	30.50 (0.334)	**1.000**	**0.985**
1000	PSBC-GL-G	60.10 (0.295)	**42.70** (0.301)	0.675	0.642
	PSBC-GL-B	63.00 (0.286)	42.30 (0.302)	0.709	0.661
	PSBC-EN-G	**80.80** (0.311)	30.20 (0.333)	**1.000**	**0.979**
	PSBC-EN-B	79.00 (0.311)	29.80 (0.333)	0.997	0.972

Figure 1 presents the overall performance of the `PSBC` models, using the corresponding BIC values of each model. The plot indicate that `PSBC-GL-B` outperformed the competing `PSBC` models, followed by `PSBC-GL-G`, and so on.

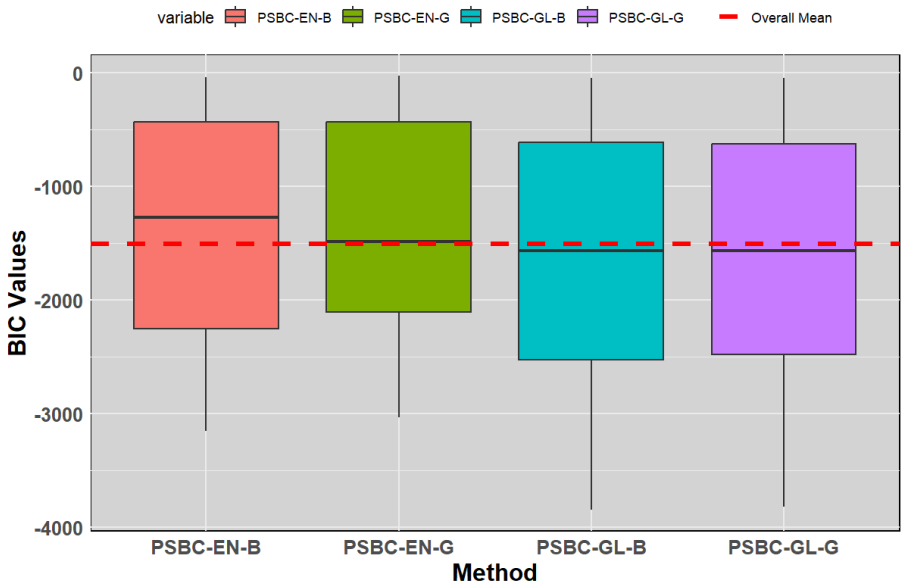

Figure 1. Overall comparison of the PSBC models on the simulated dataset.

3.2. Real-Life Studies

This section focuses on the practical application of proposed PSBC models along with existing models using three (3) different microarray gene expression datasets. The three datasets are NCI breast cancer data [45] (Data 1), Dutch breast cancer data [46] (Data 2), and Diffuse large-B-cell lymphoma data [47] (Data 3).

A detailed description of the datasets are presented in Table 5. For the first dataset, the number of observations (n) is less than that of predictors ($p < n$), while the remaining two (2) datasets include more predictors than observations ($p > n$).

The results from Table 5 reveal the dimensions of the three datasets. The first dataset (Data 1: `NKI70`) consists of low-dimensional data, since the number of individual genes is fewer than the number of individuals in the dataset. On the other hand, the remaining two datasets are high-dimensional.

Table 5. The description of the three datasets used in this study.

Data ID	Description	No. of Subjects n	No. of Predictors p
Data 1. NK170	NCI breast cancer data, [45]	144	70
Data 2. DutchBC	Dutch breast cancer data, [46]	295	4919
Data 3. DLBCL	Diffuse large-B-cell lymphoma data from patients undergoing R-CHOP treatment, [47]	181	3835

3.3. Data Description Using Kaplan–Meier Survival Curve

To assess the statistical significance of differences between treatment subgroups in the datasets, this study employed the Kaplan–Meier survival curve. This curve was utilized to analyze the duration from the manifestation of cancer symptoms to the incidence of the primary endpoint.

Figure 2 (left to right) indicates a significant difference in survival times for the patients in low- and high-risk groups (`log rank test p = 0.018`). The Kaplan–Meier survival probability estimates at 12 months were about 0.76 for the patients in the low-risk group, and about 0.375 for the patients at high risk.

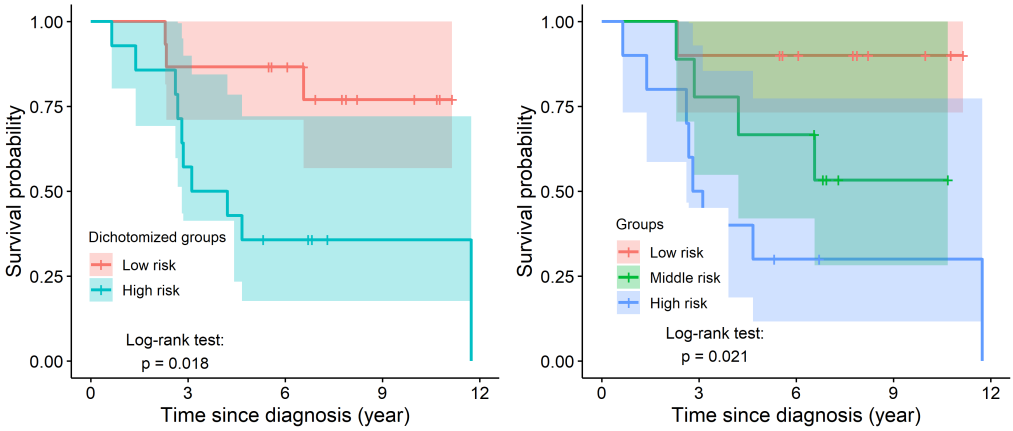

Figure 2. Kaplan–Meier survival curves depicting the duration from the onset of breast cancer symptoms to the occurrence of the primary endpoint (death) for different groups. The curves also highlight the ongoing risk of reaching the primary endpoint at various time points, indicating the number of patients still susceptible to the event [Data 1].

Also in the case where the patients were subgrouped into three groups (`Low, Middle`, and `High risk`), survival differences were noticed between the three (3) subgroups (log rank test $p = 0.021$). Additionally, the result revealed that the survival probability estimates at 12 months were well above 0.75 for patients at low risk, a little above 0.5 for patient at middle risk, and roughly above 0.25 for patients at high risk.

Figure 3 (left to eight) indicates a significant difference in survival times for the patients with breast cancer in the low- and high-risk groups (`log rank test p = 0.0094`). The Kaplan–Meier survival probability estimates at 12 months were about 0.80 for the patients in the low-risk group, and about 0.60 for the patients at high risk.

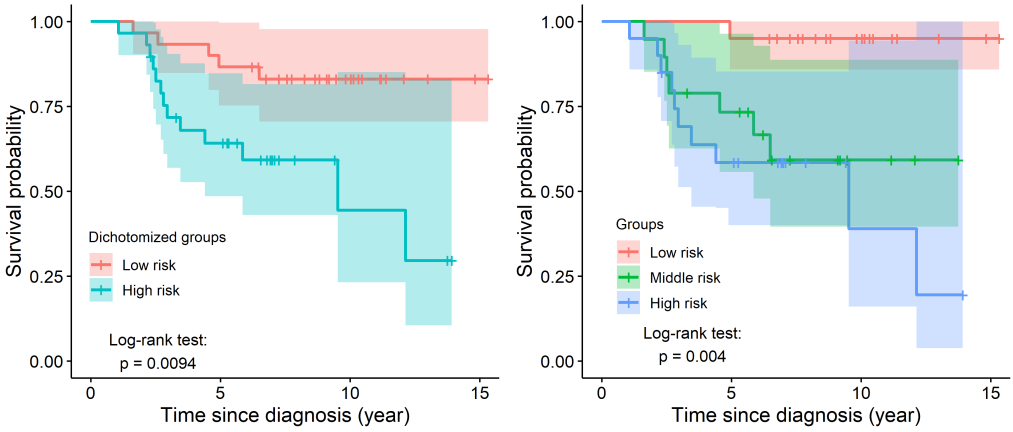

Figure 3. Kaplan–Meier survival curves depicting the duration from the onset of breast cancer symptoms to the occurrence of the primary endpoint (death) for different groups. The curves also highlight the ongoing risk of reaching the primary endpoint at various time points, indicating the number of patients still susceptible to the event [Data 2].

Furthermore, in the case where the patients were subgrouped into three groups (`Low`, `Middle`, and `High risk`), survival differences were noticed between the three (3) subgroups (log rank test $p = 0.004$). In addition, the survival probability estimates at 12 months were well above 0.90 for patients at low risk, a little above 0.80 for patient at middle risk, and roughly above 0.30 for patients at high risk.

Figure 4 (left to right) indicates a significant difference in survival times for the patients receiving rituximab immunotherapy in the low- and high-risk groups (log rank test $p = 0.63$). The Kaplan–Meier survival probability estimates at 12 months were about 0.80 for the patients in the low-risk group, and about 0.60 for the patients at high risk.

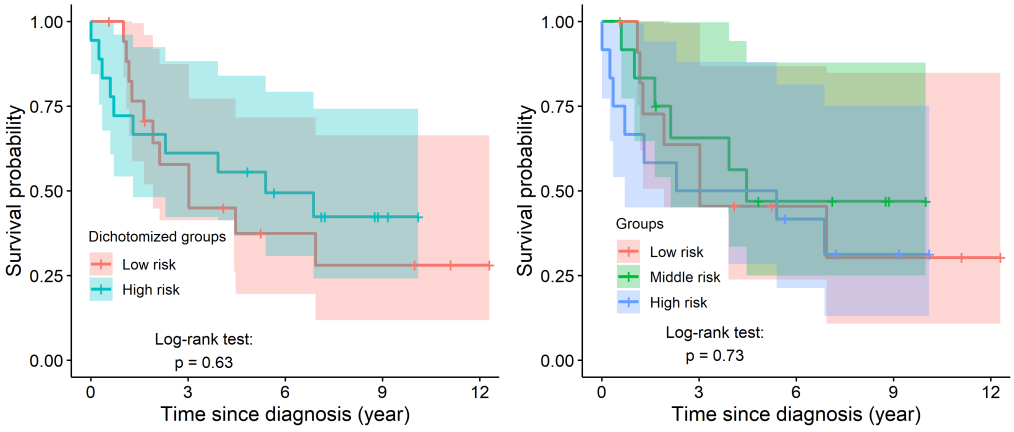

Figure 4. Kaplan–Meier survival curves depicting the duration from the onset of rituximab immunotherapy in addition to chemotherapy symptoms to the occurrence of the primary endpoint (death) for different groups. The curves also highlight the ongoing risk of reaching the primary endpoint at various time points, indicating the number of patients still susceptible to the event [Data 3].

In addition, in the case where the patients were subgrouped into three groups (`Low`, `Middle`, and `High risk`), survival differences were noticed between the three (3) subgroups (log rank test $p = 0.73$). And the survival probability estimates at 12 months were well above 0.90 for patients at low risk, a little above 0.80 for patient at middle risk, and roughly above 0.30 for patients at high risk.

3.4. Comparison of the Methods' Performance on the Real-Life Dataset

This study evaluates the performance of both the existing and proposed methods by employing BIC thresholding (Equation (15)), which facilitates effective grouped variable selection. Table 6 showcases the comparative results for both the existing and proposed methodologies.

Table 6. Real-life results based on 100 replications.

Datasets	Prior	Method	No. of Predictors Selected	BIC
Data 1	Gamma	PSBC-GL-G	31	−195.1181
		PSBC-EN-G	31	−196.9853
	Beta	PSBC-GL-B	33	−219.3794
		PSBC-EN-B	39	−168.7433
Data 2	Gamma	PSBC-GL-G	20	−26,773.67
		PSBC-EN-G	20	−26,242.25
	Beta	PSBC-GL-B	9	−26,871.26
		PSBC-EN-B	15	−26,285.90
Data 3	Gamma	PSBC-GL-G	33	−17,512.08
		PSBC-EN-G	106	−12,293.08
	Beta	PSBC-GL-B	20	−17,879.20
		PSBC-EN-B	40	−13,735.29

Similar to the results from the simulated study, Table 6 indicates that `PSBC-GL-B` demonstrates superior variable selection capability compared to the other three competing methods, as it consistently exhibits the lowest BIC values across the three datasets.

For instance, in the case of `Data 1`, `PSBC-GL-B` achieves the lowest BIC value of **−219.3794**, followed by `PSBC-EN-G`, with `PSBC-EN-B` performing the least effectively. Similarly, for `Data 2`, `PSBC-GL-B` outperforms the other methods with a BIC of **−26871.26**, followed by `PSBC-GL-G` with **−26773.67**. The trend continues with `Data 3`, where `PSBC-GL-B` again achieves the lowest BIC value of **−17879.20**.

We further our investigation by considering the performance metrics of the real-life datasets, alongside the results of the BIC and the number of variables selected. The results of these metrics, which measure the model's accuracy, are presented in Table 7. The results in Table 7 reveal competitive performance among the four methods based on the true positive eate (TPR) and true negative rate (TNR), while the existing models show slightly higher performance based on the average AUC and C-index. On the other hand, the NKI70 and DutchBC datasets demonstrate better performance across all methods compared to the DLBCL dataset. The DLBCL dataset seems to be a challenging one for all the methods, with a reduced TPR, AUC, and C-Index.We are not surprised by these prediction performances of the DLBCL dataset, as similar prediction performances were recorded in the study by Lee et al., 2015 [36].

Moreover, to account for the predictive performance of all the models over time, we visualize the AUC results in the Figure 5, since the AUC is the area under the ROC curve. The results from Figure 5 present the average AUC curves for the four methods across the three real-life datasets. Again, the results reveal competitive performance at some time points and slightly higher performance in PSBC-GL-G and PSBC-GL-B than in the other models in the DutchBC dataset (Figure 5B).

Table 7. Performance metrics of different methods across three datasets (NKI70, DutchBC, and DLBCL).

Data	Method	TPR % (SD)	TNR % (SD)	AUC	C-INDEX
NKI70	PSBC-GL-G	78.10 (0.266)	**39.30** (0.321)	0.892	0.861
	PSBC-GL-B	78.20 (0.268)	39.10 (0.321)	0.895	0.863
	PSBC-EN-G	**79.80** (0.307)	30.80 (0.334)	**0.995**	**0.971**
	PSBC-EN-B	**79.80** (0.306)	30.90 (0.334)	0.994	0.969
DutchBC	PSBC-GL-G	65.20 (0.264)	45.80 (0.302)	0.695	0.656
	PSBC-GL-B	63.70 (0.281)	**46.10** (0.295)	0.667	0.660
	PSBC-EN-G	**83.50** (0.269)	39.90 (0.319)	**0.937**	**0.913**
	PSBC-EN-B	83.10 (0.268)	40.10 (0.319)	0.932	0.910
DLBCL	PSBC-GL-G	**60.30** (0.296)	43.30 (0.294)	**0.672**	**0.628**
	PSBC-GL-B	54.70 (0.302)	45.80 (0.296)	0.589	0.571
	PSBC-EN-G	48.10 (0.309)	**51.90** (0.282)	0.461	0.462
	PSBC-EN-B	47.00 (0.302)	**51.90** (0.284)	0.451	0.451

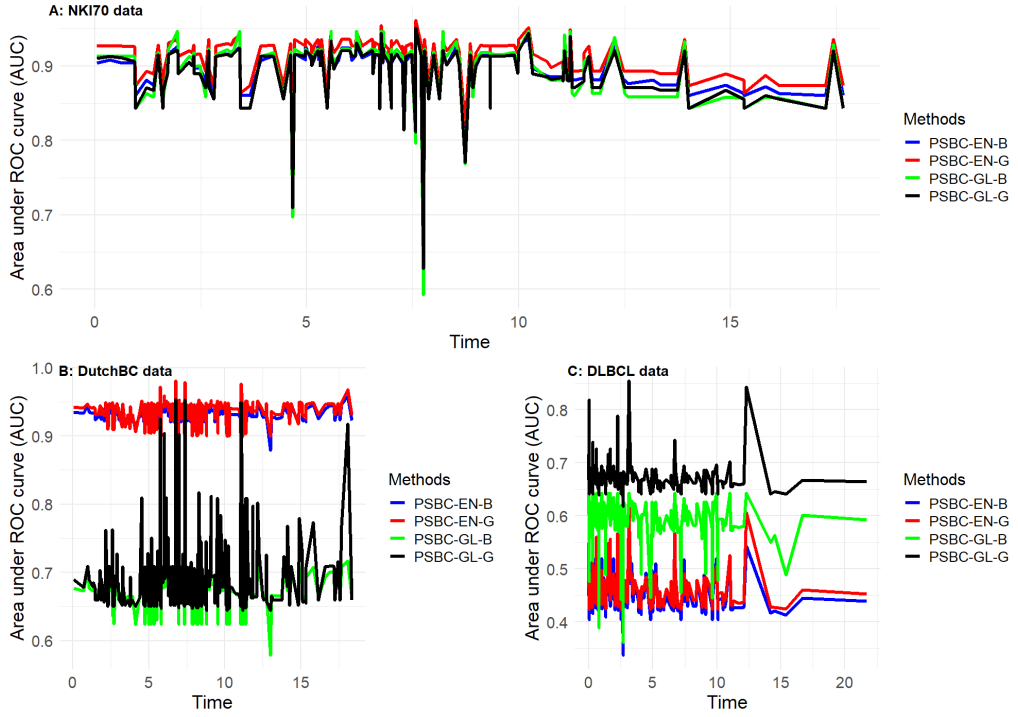

Figure 5. Prediction capability of the four models on three real-life datasets.

3.5. Posterior Credible Region

In addition to the results in the previous section on the performance of the methods using the real-life dataset, we estimated the credible interval or region for the posterior estimates using both the proposed and the competing methods. However, the solid dots denote the posterior mode of the coefficients and the lines denote the 95% confidence intervals. The closer the posterior mode is to the right side indicates significant features at a $p < 0.05$. Hence, the plot reveals that few of the features are significantly related to the survival curve.

Figures 6 and 7 depict the posterior credible region for the methods for Data 1. From the figures, we notice that in the case of `PSBC-GL-B` and `PSBC-EN-B` the dots drift more to the right-hand side compared to those of the competing methods. This simply implies that the proposed methods are more adequate in more variables in a dataset with a grouping effect.

Similar to the findings on Data 1, Figures 8 and 9 represent the posterior credible region for the methods for Data 2; this study provides the posterior mode of the coefficient estimates across the four methods. The results reveal that the proposed methods identified more significant features compared to existing methods. This is evident as the posterior-mode dots for the proposed methods drift further to the right or left (i.e., farther from the center) than those of the existing methods. In contrast, the posterior-mode dots closer to the center (0) indicate less significant features. We also observe a similar pattern in the third dataset (see Figures 10 and 11) .

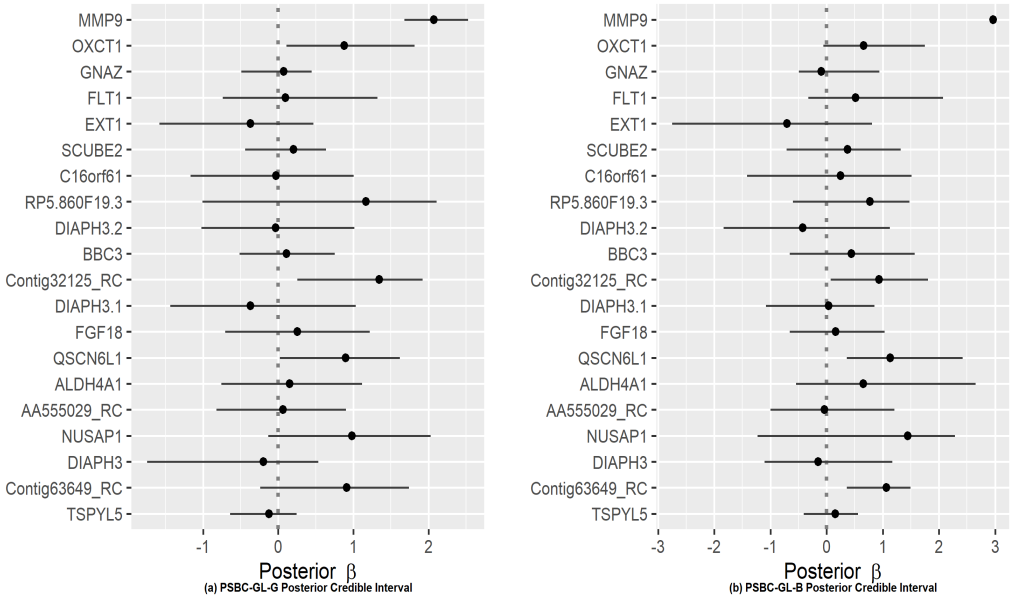

Figure 6. Posterior credible region of the 20 genetic features; only these 20 features are presented to maintain the figure's legibility for PSBC-GL: (**a**) gamma prior; (**b**) beta prior for Data 1.

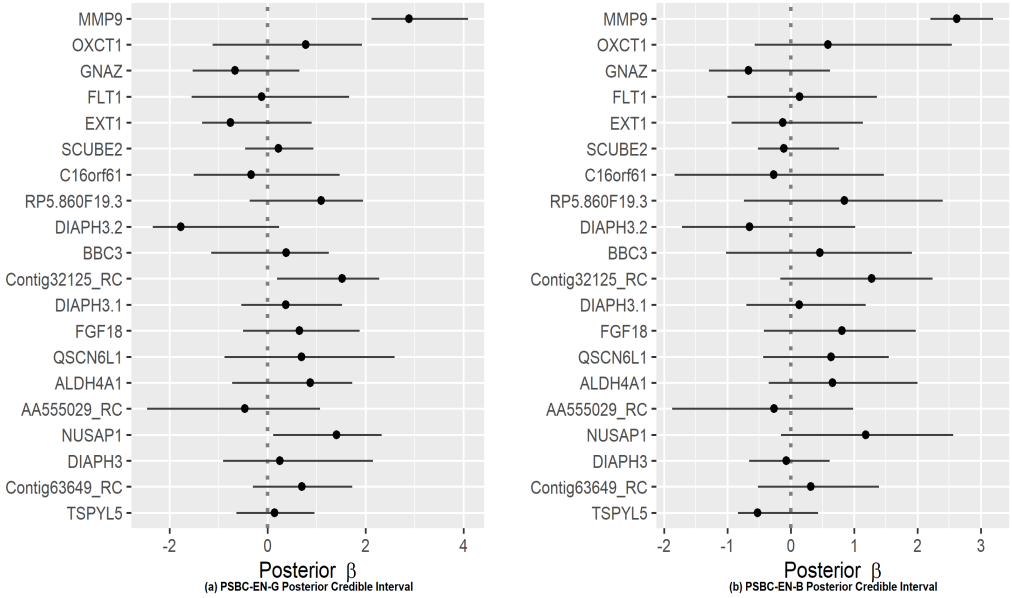

Figure 7. Posterior credible region of the 20 genetic features; only these 20 features are presented to maintain the figure's legibility for PSBC-EN: (**a**) gamma prior; (**b**) beta prior for Data 1.

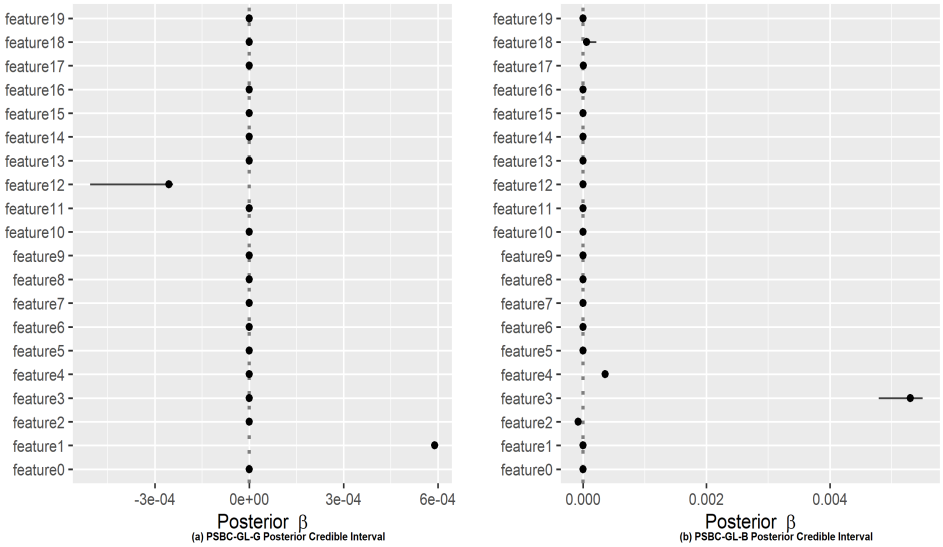

Figure 8. Posterior credible region of the 20 genetic features; only these 20 features are presented to maintain the figure's legibility for PSBC-GL: (**a**) gamma prior; (**b**) beta prior for Data 2.

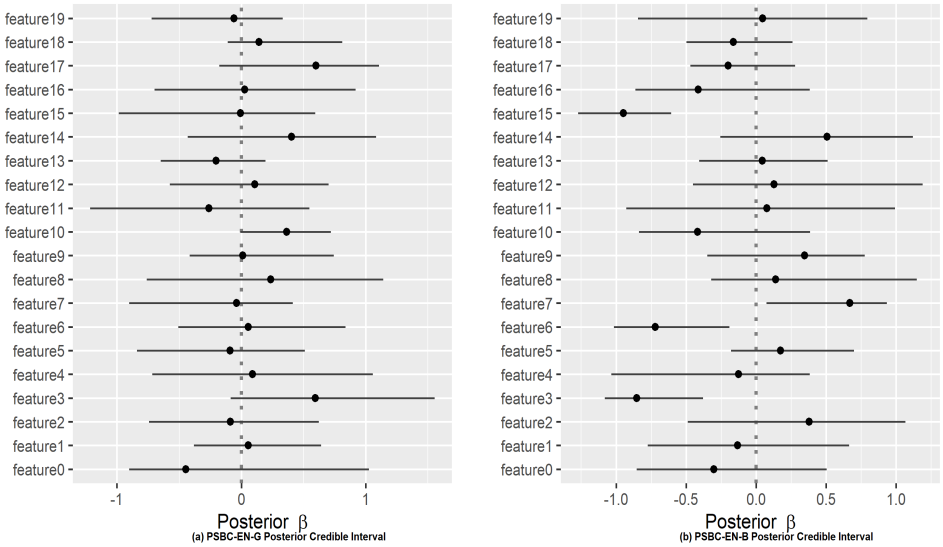

Figure 9. Posterior credible region of the 20 genetic features; only these 20 features are presented to maintain the figure's legibility for PSBC-EN: (**a**) gamma prior; (**b**) beta prior for Data 2.

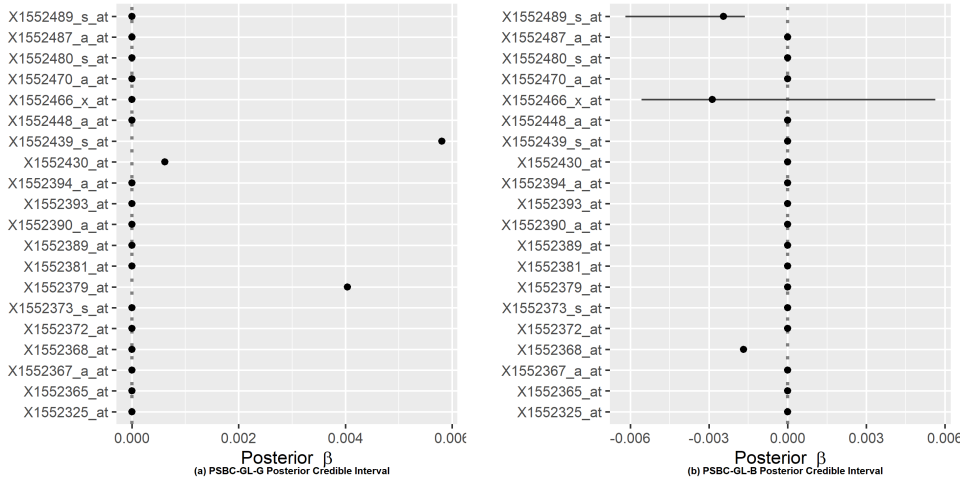

Figure 10. Posterior credible region of the 20 genetic features; only these 20 features are presented to maintain the figure's legibility for PSBC-GL: (**a**) gamma prior; (**b**) beta prior for Data 3.

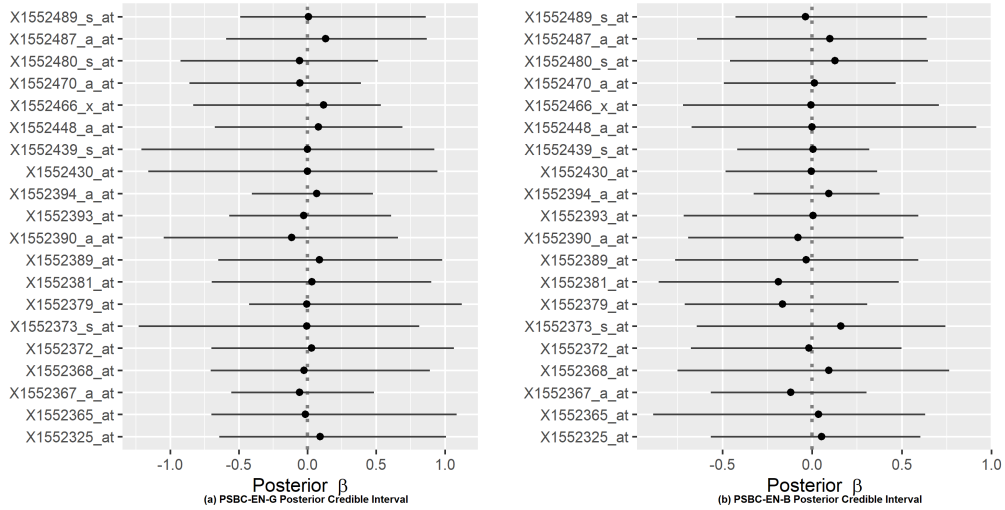

Figure 11. Posterior credible region of the 20 genetic features; only these 20 features are presented to maintain the figure's legibility for PSBC-EN: (**a**) gamma prior; (**b**) beta prior for Data 3.

4. Conclusions

In this study, we have successfully developed two new penalized Bayesian models for fast variable selection in survival analysis. The new methods (PSBC-GL-B and PSBC-EN-B) are highly efficient in identifying significant covariates, shrinking the coefficients of related (or grouped) variables toward a common value, and, as a result, uncovering any grouping behavior among the covariates, which is a common issue in large-scale omics data. Our methodologies implemented a beta process for the cumulative baseline hazard function, contrasting with the gamma process considered in the previous study, thus leading to improvements in the penalized Bayesian model framework.

Through extensive comparative methodologies, the efficacy of our proposed approaches is demonstrated. The results show that the proposed method (PSBC-GL-B) performed better under our simulation settings and real-life datasets in term of variable selection capability and prediction accuracy than the three competing methods, PSBC-GL-G, PSBC-EN-B, and PSBC-EN-G. However, in terms of performance metrics, all four models seem to compete favorably with each other. Although PSBC-GL-G shows slight improvement on real-life data compared to the other methods, PSBC-EN-B also demonstrates a slight performance advantage over others in the simulation scenarios based on the average AUC and c-index.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/computation13020021/s1.

Author Contributions: Conceptualization, K.A.D.; methodology, E.J.A. and K.A.D.; writing—original draft preparation, K.A.D.; writing E.J.A. and K.A.D.; supervision, R.K.L. and O.T.W.; software, E.J.A. and K.A.D.; formal analysis, E.J.A. and R.K.L.; validation, O.T.W. All authors have read and agreed to the published version of the manuscript.

Funding: Kazeem A. Dauda was supported by the Trond Mohn Foundation (project HyperEvol under grant agreement no. TMS2021TMT09), through the Centre for Antimicrobial Resistance in Western Norway (CAMRIA) (TMS2020TMT11).

Data Availability Statement: All data used are available in published sources.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PSBC	Penalized semi-parametric Bayesian Cox
PSBC-EN	Penalized semi-parametric Bayesian Cox with elastic-net
PSBC-GL	Penalized semi-parametric Bayesian Cox with group lasso
PSBC-EN-B	Penalized semi-parametric Bayesian Cox model with elastic-net and a beta prior
PSBC-GL-B	Penalized semi-parametric Bayesian Cox model with group lasso and a beta prior
PSBC-EN-G	Penalized semi-parametric Bayesian Cox model with elastic-net and a gamma prior
PSBC-GL-G	Penalized semi-parametric Bayesian Cox model with group lasso and a gamma prior
BIC	Bayesian information criteria
Cox-PH	Cox proportional hazard
MCMC	Markov Chain Monte Carlo
AR	Average Rank

References

1. Dauda, K.A. Optimal tuning of Random Survival Forest hyperparameter with an application to liver disease. *Malays. J. Med. Sci.* **2022**, *29*, 67–76. [CrossRef] [PubMed]
2. Ryan, W.; Ahn, M.; Yang, H. Spike-and-slab type variable selection in the Cox proportional hazards model for high-dimensional features. *J. Appl. Stat.* **2022**, *49*, 2189–2207. [CrossRef] [PubMed]
3. Paolucci, I.; Lin, Y.M.; Silva, J.; Brock, K.; Odisio, B. Bayesian parametric models for survival prediction in medical applications. *BMC Med. Res. Methodol.* **2023**, *23*, 250. [CrossRef] [PubMed]
4. Widłak, W. High-throughput technologies in molecular biology. In *Molecular biology: Not only for bioinformaticians*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 139–153.
5. Zeng, C.; Huang, J.; Wang, H.; Xie, J.; Zhang, Y. Deep Bayesian survival analysis of rail useful lifetime. *Eng. Struct.* **2023**, *295*, 116822. [CrossRef]
6. Haredasht, F.N.; Dauda, K.A.; Vens, C. Exploiting Censored Information in Self-Training for Time-to-Event Prediction. *IEEE Access* **2023**, *11*, 96831–96840. [CrossRef]
7. Singh, R.; Mukhopadhyay, K. Survival analysis in clinical trials: Basics and must know areas. *Perspect. Clin. Res.* **2011**, *2*, 145–148. [CrossRef]
8. Dauda, K.A.; Pradhan, B.; Shankar, B.U.; Mitra, S. Decision tree for modeling survival data with competing risks. *Biocybern. Biomed. Eng.* **2019**, *39*, 697–708. [CrossRef]
9. Dauda, K.A.; Yahya, W.B.; Banjoko, A.W.; Olorede, K.O. Competing Risk Modeling Using Cumulative Incidence Function: Application to Recurrent Bladder Cancer data. *FUOYE J. Eng. Technol. Fed. Univ. Sci.* **2018**, *3*, 127–132.
10. Dauda, K.A.; Lamidi, R.K.; Dauda, A.A.; Yahya, W.B. A New Generalized Gamma-Weibull Distribution with Applications to Time-to-event Data. *bioRxiv* **2023**. [CrossRef]
11. Zhao, L. Deep neural networks for predicting restricted mean survival times. *Bioinformatics* **2020**, *36*, 5672–5677. [CrossRef]
12. Lee, C.; Zame, W.; Yoon, J.; Van Der Schaar, M. Deephit: A deep learning approach to survival analysis with competing risks. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32, pp. 2314–2321.
13. Dauda, K.A.; Yahya, W.B.; Banjoko, A.W. Survival analysis with multivariate adaptive regression splines using Cox-Snell residual. *Ann. Comput. Sci. Ser.* **2015**, *13*, 25–41.
14. Feng, D.; Zhao, L. BDNNSurv: Bayesian Deep Neural Networks for Survival Analysis Using Pseudo Values. *J. Data Sci.* **2021**, *19*, 542–554. [CrossRef]
15. Tang, Z.; Shen, Y.; Zhang, X.; Yi, N. The spike-and-slab lasso Cox model for survival prediction and associated genes detection. *Bioinformatics* **2017**, *33*, 2799–2807. [CrossRef] [PubMed]
16. Banerjee, S.; Castillo, I.; Ghosal, S. Bayesian inference in high-dimensional models. *arXiv* **2021**, arXiv:2101.04491.
17. Samorodnitsky, S.; Hoadley, K.; Lock, E. A hierarchical spike-and-slab model for pan-cancer survival using pan-omic data. *BMC Bioinform.* **2022**, *23*, 235. [CrossRef]
18. Ishwaran, H.; Rao, J.S. Spike and slab variable selection: Frequentist and Bayesian strategies. *Ann. Stat.* **2005**, *33*, 730. [CrossRef]
19. Yang, H.; Baladandayuthapani, V.; Rao, A.U.; Morris, J.S. Quantile function on scalar regression analysis for distributional data. *J. Am. Stat. Assoc.* **2019**, *115*, 90–106. [CrossRef]

20. Maity, A.K.; Bhattacharya, A.; Mallick, B.K.; Baladandayuthapani, V. Bayesian data integration and variable selection for pan-cancer survival prediction using protein expression data. *Biometrics* **2020**, *76*, 316–325. [CrossRef]
21. Mao, Y.; Wang, L.; Lin, X.; Sui, X. Bayesian variable selection in joint modeling of longitudinal data and interval-censored failure time data. *Res. Sq.* **2024** [CrossRef]
22. Cox, D.R. Regression models and life-tables. *J. R. Stat. Soc. Ser. Methodol.* **1972**, *34*, 187–202. [CrossRef]
23. Tachmazidou, I.; Andrew, T.; Verzilli, C.J.; Johnson, M.R.; De Iorio, M. Bayesian survival analysis in genetic association studies. *Bioinformatics* **2008**, *24*, 2030–2036. [CrossRef] [PubMed]
24. Kalbfleisch, J.D. Non-parametric Bayesian analysis of survival time data. *J. R. Stat. Soc. Ser. Methodol.* **1978**, *40*, 214–221. [CrossRef]
25. Ibrahim, J.G.; Chen, M.H.; MacEachern, S.N. Bayesian variable selection for proportional hazards models. *Can. J. Stat.* **1999**, *27*, 701–717. [CrossRef]
26. Kyung, M.; Gill, J.; Ghosh, M.; Casella, G. Penalized Regression, Standard Errors, and Bayesian Lassos. *Bayesian Anal.* **2010**, *5*, 369–412.
27. Ouyang, B.; Sinha, D.; Slate, E.H.; Van Bakel, A.B. Bayesian analysis of recurrent event with dependent termination: An application to a heart transplant study. *Stat. Med.* **2013**, *32*, 2629–2642. [CrossRef]
28. Sinha, A.; Chi, Z.; Chen, M.H. Bayesian inference of hidden gamma wear process model for survival data with ties. *Stat. Sin.* **2015**, *25*, 1613. [CrossRef]
29. Li, Y.; Seo, S.; Lee, K.H. Bayesian survival analysis using gamma processes with adaptive time partition. *J. Stat. Comput. Simul.* **2021**, *91*, 2937–2952. [CrossRef]
30. Gu, X.; Yin, G.; Lee, J.J. Bayesian two-step Lasso strategy for biomarker selection in personalized medicine development for time-to-event endpoints. *Contemp. Clin. Trials* **2013**, *36*, 642–650. [CrossRef]
31. Zhao, L.; Feng, D.; Bellile, E.L.; Taylor, J.M. Bayesian random threshold estimation in a Cox proportional hazards cure model. *Stat. Med.* **2014**, *33*, 650–661. [CrossRef]
32. Katzman, J.L.; Shaham, U.; Cloninger, A.; Bates, J.; Jiang, T.; Kluger, Y. DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med. Res. Methodol.* **2018**, *18*, 24. [CrossRef]
33. Dauda, K.A.; Olorede, K.O.; Banjoko, A.W.; Yahya, W.B.; Ayipo, Y.O. Genetic Diagnosis, Classification, and Risk Prediction in Cancer Using Next-Generation Sequencing in Oncology. In *Computational Approaches in Biomaterials and Biomedical Engineering Applications*; CRC Press: Boca Raton, FL, USA, 2024; pp. 107–122.
34. Lee, K.H.; Chakraborty, S.; Sun, J. Bayesian variable selection in semiparametric proportional hazards model for high dimensional survival data. *Int. J. Biostat.* **2011**, *7*, 00001022021557467913O1. [CrossRef]
35. Lee, K.H.; Chakraborty, S.; Sun, J. Variable selection for high-dimensional genomic data with censored outcomes using group lasso prior. *Comput. Stat. Data Anal.* **2017**, *112*, 1–13. [CrossRef]
36. Lee, K.H.; Chakraborty, S.; Sun, J. Survival prediction and variable selection with simultaneous shrinkage and grouping priors. *Stat. Anal. Data Mining Asa Data Sci. J.* **2015**, *8*, 114–127. [CrossRef]
37. Anirban, R.; Brian, K. Gamma Processes, Stick-Breaking, and Variational Inference. *Artif. Intell. Stat.* **2015**, *38*, 800–808.
38. Giordano, R.; Liu, R.; Jordan, M.I.; Broderick, T. Evaluating Sensitivity to the Stick-Breaking Prior in Bayesian Nonparametrics (with Discussion). *Bayesian Anal.* **2023**, *18*, 287–366. [CrossRef]
39. Paisley, J.; Carin, L. Nonparametric factor analysis with beta process priors. In Proceedings of the International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009.
40. Adeniyi, E.J. Bayesian Survival Analysis with Flexible Penalization Using Beta Process Prior for Baseline Hazard. Master's Thesis, Kwara State University, Malete, Nigeria, 2023.
41. Lee, K.H. *Bayesian Variable Selection in Parametric and Semiparametric High Dimensional Survival Analysis*; University of Missouri-Columbia: Columbia, MO, USA, 2011.
42. Kamarudin, A.N.; Cox, T.; Kolamunnage-Dona, R. Time-dependent ROC curve analysis in medical research: Current methods and applications. *BMC Med. Res. Methodol.* **2017**, *17*, 53. [CrossRef]
43. Heagerty, P.J.; Lumley, T.; Pepe, M.S. Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics* **2000**, *56*, 337–344. [CrossRef]
44. Harrell, F.E.; Califf, R.M.; Pryor, D.B.; Lee, K.L.; Rosati, R.A. Evaluating the yield of medical tests. *JAMA* **1982**, *247*, 2543–2546. [CrossRef]
45. Van't Veer, L.J.; Dai, H.; Van De Vijver, M.J.; He, Y.D.; Hart, A.A.; Mao, M.; Peterse, H.L.; Van Der Kooy, K.; Marton, M.J.; Witteveen, A.T.; et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **2002**, *415*, 530–536. [CrossRef]

46. Van Houwelingen, H.C.; Bruinsma, T.; Hart, A.A.; Van't Veer, L.J.; Wessels, L.F. Cross-validated Cox regression on microarray gene expression data. *Stat. Med.* **2006**, *25*, 3201–3216. [CrossRef]
47. Lenz, G.; Wright, G.; Dave, S.; Xiao, W.; Powell, J.; Zhao, H.; Xu, W.; Tan, B.; Goldschmidt, N.; Iqbal, J.; et al. Stromal gene signatures in large-B-cell lymphomas. *N. Engl. J. Med.* **2008**, *359*, 2313–2323. [CrossRef] [PubMed]