mathematics

MDPI

Article

Semiparametric Analysis of Additive–Multiplicative Hazards Model with Interval-Censored Data and Panel Count Data

Tong Wang [1], Yang Li [1], Jianguo Sun [2] and Shuying Wang [1,*]

[1] School of Mathematics and Statistics, Changchun University of Technology, Changchun 130012, China; wangtong0783@163.com (T.W.); 17519403278@163.com (Y.L.)
[2] Department of Statistics, University of Missouri, Columbia, MO 65211, USA; sunj@missouri.edu
* Correspondence: wangshuying0601@163.com

Abstract: In survival analysis, interval-censored data and panel count data represent two prevalent types of incomplete data. Given that, within certain research contexts, the events of interest may simultaneously involve both data types, it is imperative to perform a joint analysis of these data to fully comprehend the occurrence process of the events being studied. In this paper, a novel semiparametric joint regression analysis framework is proposed for the analysis of interval-censored data and panel count data. It is hypothesized that the failure time follows an additive–multiplicative hazards model, while the recurrent events follow a nonhomogeneous Poisson process. Additionally, a gamma-distributed frailty is introduced to describe the correlation between the failure time and the count process of recurrent events. To estimate the model parameters, a sieve maximum likelihood estimation method based on Bernstein polynomials is proposed. The performance of this estimation method under finite sample conditions is evaluated through a series of simulation studies, and an empirical study is illustrated.

Keywords: interval-censoring; frailty model; additive–multiplicative hazards model; sieve maximum likelihood estimation

MSC: 62N02

Citation: Wang, T.; Li, Y.; Sun, J.; Wang, S. Semiparametric Analysis of Additive–Multiplicative Hazards Model with Interval-Censored Data and Panel Count Data. *Mathematics* **2024**, *12*, 3667. https://doi.org/10.3390/math12233667

Academic Editor: Li-pang Chen

Received: 4 November 2024
Revised: 20 November 2024
Accepted: 21 November 2024
Published: 22 November 2024

1. Introduction

In survival analysis, researchers typically focus on the occurrence time of events of interest. However, in certain situations, due to various factors, it is not possible to precisely observe the time at which failure events occur, and it can only be determined that they occur within some time interval. The survival data obtained under such circumstances are referred to as interval-censored data. On the other hand, when the focus of study is on recurrent events, the research is concerned not only with the time points of the event occurrence but also with the counting process of these events. Since the continuous monitoring of subjects is usually not feasible and observations can only be made at specific discrete time points, the resulting data type is panel count data. Interval-censored data and panel count data are two common types of incomplete data, and a variety of models and methods have been established for the analysis of these data types. Zhu et al. [1], Sun and Ding [2] and Bouaziz et al. [3] engage in modeling research on interval-censored data utilizing a variety of models. For the panel count data, relevant developments include the works of Sun and Kalbfleisch [4], Wellner and Zhang [5], Hu et al. [6], Hua and Zhang [7], Hua et al. [8] and Yao et al. [9]. In particular, in the study conducted by Wellner and Zhang [10], a semiparametric regression model based on the nonhomogeneous Poisson process was investigated, with a detailed analysis of the mean of the panel count process. Subsequently, the Poisson process has emerged as a significant tool for analyzing such data.

However, in multiple disciplines, including medical research and social sciences, the coexistence of interval-censored data and panel count data is a prevalent phenomenon.

For instance, in clinical research, multiple follow-up time points are typically established. Between these time points, researchers aim to observe the occurrence of specific events and their potential recurrence. Due to this observational approach, the resulting data inherently encompass both interval-censored and panel count data. Consequently, conducting a joint modeling analysis of these two types of data is particularly crucial. Wen and Chen have proposed various methodologies for analyzing current count and current status data. Specifically, they initially introduced a full likelihood approach in reference [11], which conducts a joint analysis of these data with a gamma-distributed frailty. Subsequently, in reference [12], they relaxed the assumptions regarding the frailty and proposed a pseudo-sufficient likelihood method. To more extensively analyze general interval-censored data that include recurrent and nonrecurrent events, they proposed a novel modeling framework in reference [13], which no longer relies on the assumption of a frailty distribution.

In another study, Xu et al. [14] introduced a full likelihood method for the integrated analysis of panel count and interval-censored data, which is characterized by a shared gamma frailty influencing both events. It should be noted that Xu et al. [14] applied the Cox proportional hazards model to analyze the nonrecurrent event. And the proportional hazards assumption is not always met in practical scenarios. An alternative and widely applied model is the additive hazards model, where the covariate effects are additive. However, in real-world problems, the impact of covariates on the failure time may not be limited to multiplicative or additive effects. Lin and Ying [15] also proposed the additive–multiplicative hazards model, which allows some covariates to have multiplicative effects, other covariates to have additive effects, or allows covariates to have additive and multiplicative effects simultaneously. The proportional hazards model and the additive hazards model are both special cases of this model, thereby endowing the additive–multiplicative hazards model with a more robust modeling capability.

In this article, we concentrate on a joint approach to analyzing interval-censored and panel count data, employing a frailty variable to characterize the linkage between the failure time process and recurrent event process. Additionally, we apply an additive–multiplicative hazards model, inclusive of a shared frailty, to model the failure times. Furthermore, since the additive–multiplicative hazards model is a semiparametric model, it contains an infinite-dimensional nonparametric part of the unknown baseline function and a finite-dimensional parametric part of interest, both of which need to be estimated simultaneously. Therefore, we use the sieve maximum likelihood estimation (SMLE) method. SMLE approximates the infinite-dimensional parameter space with a finite-dimensional parameter space, simplifying the estimation problem, and the number of unknown parameters used to define the approximation space increases slowly with the sample size. It is this feature that makes the sieve method surpass the classical parametric methods that use a fixed finite-dimensional parameter space, offering better flexibility and robustness [16]. Additionally, we use Bernstein polynomials to approximate the unknown baseline function; the Bernstein polynomial possesses an optimal shape-preserving attribute when compared to other approximating polynomials [17], and it offers a simpler application process compared to spline approximations, as it eliminates the need to define interior knots. Thus, we suggest employing a sieve maximum likelihood method, which is based on Bernstein polynomials, for the estimation of the model parameters.

The subsequent content of this paper is structured as follows: Section 2 provides a detailed exposition of the data and models employed in this research. Section 3 details the derivation of the likelihood function and introduces the sieve maximum likelihood estimation procedure that is proposed within this work. Section 4 encompasses a series of simulation studies designed to evaluate the performance of the proposed estimation method with finite samples. Section 5 showcases the application of the proposed approach to a dataset from a skin cancer chemoprevention trial. Finally, Section 6 provides a concluding discussion.

2. Data and Model

Focusing on the study of a historical event, there are n independent subjects involved who are subject to two types of events: a nonrecurrent failure event and occurrences of a recurrent event. Suppose that there exists p-dimensional and q-dimensional vectors of covariates denoted by Z_i and X_i for subject $i, i = 1, 2, \cdots, n$; let T_i represent the timing of the failure event; and let $N_i(t)$ represent the cumulative number of recurrence events that have occurred by time t. Suppose that for subject i, a sequence of observation times $s_{i1} < s_{i2} < \cdots < s_{im_i}$ is established, where m_i denotes the count of observations made on subject i. Accordingly, the information observed for each subject is limited to

$$O = \Big\{ O_i = \big(m_i, s_{ij}, N_i(s_{ij}), Z_i, X_i, \delta_{ik} = I(s_{i,k-1} < T_i \le s_{ik}), \\ j = 1, 2, \cdots, m_i, k = 1, 2, \cdots, m_i + 1\big); i = 1, 2, \cdots, n \Big\},$$

where $s_{i0} = 0$ and $s_{i,m_i+1} = \infty$. In other words, the available data consist solely of case K interval-censored data regarding the T_i and panel count data regarding the $N_i(t)$.

We introduce a latent variable η_i to model the effects of the covariates on both the time-to-event T_i and the cumulative number of recurrent events $N_i(t)$, as well as the possible correlation between T_i and $N_i(t)$, and η_i possesses a mean of 1 and a variance σ^2 that is both unknown and positive, and suppose that η_i is independent of $\{m_i, T_i, N_i(t)\}$, and given $\{Z_i, X_i, \eta_i\}$, T_i and $N_i(t)$ are also independent. Suppose that given $\{Z_i, X_i, \eta_i\}$, the hazard function of T_i has the form of

$$\lambda_i(t \mid Z_i, X_i, \eta_i) = \eta_i\big(\alpha^\top Z_i + \lambda_1(t)\exp(\beta^\top X_i)\big), \tag{1}$$

where $\lambda_1(t)$ denotes an unknown baseline hazard function, and let $\Lambda_1(t) = \int_0^t \lambda_1(u)du$ denote the baseline cumulative hazard function, α and β are vectors of unknown regression parameters. That is, T_i follows the additive–multiplicative hazards frailty model.

For the process $N_i(t)$, we consider it to be a nonhomogeneous Poisson process, with its proportional mean function being

$$E\{N_i(t) \mid X_i, \eta_i\} = \eta_i \Lambda_2(t)\exp(\gamma^\top X_i), \tag{2}$$

where $\Lambda_2(t)$ represents a baseline mean function that is unknown and nondecreasing, γ is a regression coefficient, which is a vector of dimension q. It should be noted that η_i measures the degree of association between T_i and $N_i(t)$, and it suggests that T_i and $N_i(t)$ are independent given covariates when $\eta_i = 1$. Then, we will assume that $\{m_i, s_{ij}; j = 1, 2, \cdots, m_i\}$ and $\{\eta_i, T_i, N_i(t)\}$ are independent given covariates Z_i, X_i, and the conditional distribution of $\{m_i, s_{ij}; j = 1, 2, \cdots, m_i\}$ is independent of the parameters within models (1) and (2).

3. Sieve Maximum Likelihood Approach

Define $\theta = \big(\theta^{*\top}, \Lambda^\top(\cdot)\big)^\top$ with $\theta^* = \big(\alpha^\top, \beta^\top, \gamma^\top, \sigma^2\big)^\top$ and $\Lambda(\cdot) = (\Lambda_1(\cdot), \Lambda_2(\cdot))^\top$. Subsequently, the likelihood function for θ is formulated as follows:

$$L_n(\theta; O) = \prod_{i=1}^{n} L(\theta; O_i),$$

where

$$L(\theta;O_i) = E_{\eta_i}\Bigg\{ \prod_{k=1}^{m_i+1} \Big[\exp\Big(-\eta_i\big(\Lambda_1(s_{i,k-1})e^{\beta^\top X_i} + \alpha^\top Z_i s_{i,k-1}\big)\Big) - \exp\Big(-\eta_i\big(\Lambda_1(s_{ik})e^{\beta^\top X_i} + \alpha^\top Z_i s_{ik}\big)\Big)\Big]^{\delta_{ik}} \cdot \prod_{j=1}^{m_i} \Big[\Big(\eta_i \Delta\Lambda_2(s_{ij}) \exp(\gamma^\top X_i)\Big)^{\Delta N_i(s_{ij})} \cdot \exp(-\eta_i \Delta\Lambda_2(s_{ij}) \exp(\gamma^\top X_i))\Big] \Bigg\},$$

and

$$\Delta N_i(s_{ij}) = N_i(s_{ij}) - N_i(s_{i,j-1}), \Delta N_i(s_{i1}) = N_i(s_{i1}),$$
$$\Delta\Lambda_2(s_{ij}) = \Lambda_2(s_{ij}) - \Lambda_2(s_{i,j-1}), \Delta\Lambda_2(s_{i1}) = \Lambda_2(s_{i1}).$$

Given the assumption that η_i follows a gamma distribution, the likelihood contribution for θ, denoted as $L(\theta;O_i)$, can be expressed in a simplified form:

$$L(\theta;O_i) = Q_i \prod_{k=1}^{m_i+1} (A_{i,k-1} - A_{ik})^{\delta_{ik}},$$

where

$$Q_i = \frac{\big(\sigma^2 \exp(\gamma^\top X_i)\big)^{N_i}\Gamma(\sigma^{-2}+N_i)}{\Gamma(\sigma^{-2})} \cdot \prod_{j=1}^{m_i} \big[\Delta\Lambda_2(s_{ij})^{\Delta N_i(s_{ij})}\big],$$
$$A_{i,0} = \big[1 + \sigma^2\Lambda_2(s_{i,m_i}) \exp(\gamma^\top X_i)\big]^{-\sigma^{-2}-N_i},$$
$$A_{i,k} = \big[1 + \sigma^2\Lambda_2(s_{i,m_i}) \exp(\gamma^\top X_i) + \sigma^2\Lambda_1(s_{ik}) \exp(\beta^\top X_i) + \sigma^2\alpha Z_i s_{ik}\big]^{-\sigma^{-2}-N_i},$$
$$A_{i,m_i+1} = 0$$

with $N_i = \sum_{j=1}^{m_i} \Delta N_i(s_{ij}) = N_i(s_{i,m_i})$. In the subsequent discussion, we will consider the scenario where η_i are distributed according to the gamma distribution. However, it should be noted that the method we develop is not limited to this specific case and can be generalized.

In the discussion to follow, we turn our attention to the estimation of the parameters θ. It becomes evident that the conventional approach would be the maximization of the likelihood function $L_n(\theta;O)$. However, this approach presents challenges, as the complexity arises from the necessity to estimate both the unspecified baseline hazard function $\Lambda(\cdot)$ and the unknown parameters θ^* concurrently within the model.

Thus, instead, consistent with Huang and Rossini's approach [18], we advocate for the use of sieve maximum likelihood estimation. In this method, the unknown function within the likelihood is approximated through a linear superposition of certain known basis functions, thereby constructing a sieve likelihood. Consequently, the challenge of maximizing the likelihood for the unknown function is repurposed to focus on maximizing the sieve likelihood for the coefficients within this linear superposition. This approach considerably simplifies the optimization challenge by requiring fewer basis functions to approximate the unknown function effectively; the necessary quantity of basis functions for a reasonable approximation of the unknown function increases at a considerably slower pace than the growth in the function's complexity.

More specifically, we first employ Bernstein polynomials to approximate $\Lambda_1(\cdot)$ and $\Lambda_2(\cdot)$ across the interval $[a,b]$, where a and b represent the observation time's lower and upper limits, respectively. In greater detail, define

$$\Theta = \Big\{\theta = \Big(\alpha^\top, \beta^\top, \gamma^\top, \sigma^2, \Lambda_1(\cdot), \Lambda_2(\cdot)\Big)\Big\} = \mathcal{B} \otimes \mathcal{M}^1 \otimes \mathcal{M}^2,$$

denote the parameter space of θ and define the Sieve space

$$\Theta_n = \left\{\theta_n = \left(\alpha^\top, \beta^\top, \gamma^\top, \sigma^2, \Lambda_{1n}(\cdot), \Lambda_{2n}(\cdot)\right)\right\} = \mathcal{B} \otimes \mathcal{M}_n^1 \otimes \mathcal{M}_n^2,$$

where

$$\mathcal{B} = \left\{\left(\alpha^\top, \beta^\top, \gamma^\top, \sigma^2\right) \in R^{p+2q+1}, \|\alpha\| + \|\beta\| + \|\gamma\| + |\sigma^2| \le M\right\},$$

$$\mathcal{M}_n^1 = \left\{\Lambda_{1n}(t) = \sum_{l=0}^{m} \phi_l B_l(t, m, a, b), \sum_{0 \le l \le m} |\phi_l| \le M_n, 0 \le \phi_0 \le \phi_1 \le \cdots \le \phi_m\right\},$$

$$\mathcal{M}_n^2 = \left\{\Lambda_{2n}(t) = \sum_{l=0}^{m} \xi_l B_l(t, m, a, b), \sum_{0 \le l \le m} |\xi_l| \le M_n, 0 \le \xi_0 \le \xi_1 \le \cdots \le \xi_m\right\},$$

in which ϕ_l and ξ_l are unknown parameters to be estimated and

$$B_l(t, m, a, b) = C_m^l \left(\frac{t-a}{b-a}\right)^l \left(1 - \frac{t-a}{b-a}\right)^{m-l},$$

and m denotes the degree of the Bernstein polynomials, which is usually taken to be $m = o(n^\nu)$ for some $0 < \nu < 1/2$. According to Lorentz [19], $\Theta_n = \mathcal{B} \otimes \mathcal{M}_n^1 \otimes \mathcal{M}_n^2$ can be used as the sieve space of Θ, so it is natural to define the sieve maximum likelihood estimate as $\hat{\theta}_n = \left(\hat{\alpha}_n^\top, \hat{\beta}_n^\top, \hat{\gamma}_n^\top, \hat{\sigma}_n^2, \hat{\Lambda}_{1n}(\cdot), \hat{\Lambda}_{2n}(\cdot)\right)^\top$, obtained by maximizing the log-likelihood function $l_n(\theta) = \log L_n(\theta)$ on the sieve space as the estimate for parameter θ.

Employing Bernstein polynomials transforms the estimation issue, which encompasses both finite- and infinite-dimensional parameters, into a more tractable problem that pertains solely to finite-dimensional parameters. Naturally, alternatives like splines and piecewise linear functions could be deployed for approximations. An advantage of utilizing Bernstein polynomials is their inherent capacity to model the non-negativity and monotonicity of $\Lambda_1(\cdot)$ and $\Lambda_2(\cdot)$. This can be achieved with straightforward constraints that are effortlessly addressed via reparameterization during computational processes. It has been demonstrated that the size of the sieve space described can be controlled through the expression $M_n = O(n^c)$, with c being a positive constant, as referenced in [17,20].

In the above estimation procedure, two issues need to be noted. One is that due to the non-negative and monotonic properties of functions $\Lambda_1(\cdot)$ and $\Lambda_2(\cdot)$, there are some restrictions on the parameters, which can be solved by reparameterization. In detail, a conventional method is to redefine the frailty variance parameter σ^2 in terms of its exponential form $\exp(\sigma^{2*})$. Furthermore, the parameters $\{\phi_0, \phi_1, \cdots, \phi_m\}$ and $\{\xi_0, \xi_1, \cdots, \xi_m\}$ are reparameterized by the cumulative sums of $\left\{\exp(\phi_0^*), \exp(\phi_1^*), \cdots, \exp(\phi_m^*)\right\}$ and $\left\{\exp(\xi_0^*), \exp(\xi_1^*), \cdots, \exp(\xi_m^*)\right\}$, respectively. Another issue is the selection of the Bernstein polynomial's degree for the parameter space Θ_n, as it directly affects the approximation's accuracy and level of smoothness. Apparently, a straightforward solution for this issue is to employ various values of m and subsequently compare the outcomes.

For the estimation of the covariance matrix of $\hat{\theta}_n^*$, a natural method involves utilizing the inverse of the information matrix derived from the log-likelihood function $l_n(\theta)$. Given that the likelihood of the observed data is available in a closed form, a natural estimator for the variance–covariance matrix is $I^{-1}(\hat{\theta})$, where $I(\theta)$ represents the observed information matrix, also known as the Hessian matrix, defined as $I(\theta) = -\partial^2 l_n(\theta)/\partial\theta\partial\theta^\top$. Calculating the mixed partial derivatives within $I(\theta)$ analytically can be quite complex. An alternative approach is to employ Louis's method to assess $I(\theta)$, although this method also faces similar complexities.

Therefore, we provide another simple and feasible method, namely, the profile likelihood method proposed by Murphy and Vaart [21], which is also used by Zeng et al. [22] to approximate the covariance matrix. Specifically, let e_i denote a q-dimensional vector that has the value 1 at the ith position and 0 in all other positions, and let h_n represent a

positive constant with a magnitude in the order of $n^{-1/2}$. Then, the covariance of $\hat{\theta^*}$ can be approximated by taking the inverse of a matrix where the (i,j)th element is given by

$$-\frac{pl_n(\hat{\theta_n^*}+h_ne_i+h_ne_j)-pl_n(\hat{\theta_n^*}+h_ne_i)-pl_n(\hat{\theta_n^*}+h_ne_j)+pl_n(\hat{\theta_n^*})}{h_n^2},$$

where $pl_n(\theta^*)=\max_\Lambda l_n(\theta)$ is the profile likelihood function of θ^*, which can be obtained by maximizing $\Lambda(\cdot)$ at fixed θ^*.

4. Simulation Study

We performed a comprehensive simulation study, focusing on the estimation of regression parameters, in order to evaluate the performance of the proposed estimation method with finite samples. First, the generation of latent variables η_i was based on a gamma distribution with a specified mean of 1 and a positive variance of σ^2. Subsequently, the count of observation times m_i was distributed uniformly across the options $\{1,2,3,4,5\}$. With m_i determined, the sequence of observation times $s_{i1}<s_{i2}<\cdots<s_{i,m_i}$ was taken to be the order statistics of the m_i random variables sampled from the uniform distribution ranging from 0.02 to 3. Moreover, we evaluated two situations involving covariates: (i) Z_i and X_i are one-dimensional cases and (ii) Z_i and X_i are two-dimensional cases. In these cases, Z_i was generated from a Bernoulli distribution with a probability of success of 0.5, while X_i was generated from the standard normal distribution. Then, given the Z_i, X_i and η_i, the failure times T_i were generated under model (1) and the panel count data $N_i(s_{ij})$ under model (2). The results given below are based on 1000 replications.

Table 1 presents the estimation results of parameters with their actual values from the population, specified as $\alpha=0.5$ or 0, $\beta=0.5,-0.5$ or 0, and $\sigma^2=0.5$, for the case of the one-dimensional covariates. Here, $\Lambda_1=0.1\cdot t$ and $\Lambda_2=0.5\cdot t$, and the degree of the Bernstein polynomial is determined by $m=[n^{1/4}]=3$ when the sample size $n=200$ and $m=[n^{1/4}]=4$ when the sample size $n=400$. The presented results encompass the estimated bias (Bias), defined as the mean of the point estimates minus the actual value, the sample standard errors (SSEs), the average of the estimated standard errors (ESEs), and the 95% empirical coverage probability (CP).

Table 1. Parameter estimation results under one-dimensional covariates.

Para	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
		n = 200				**n = 400**		
$\alpha=0.5$	0.0216	0.1291	0.1278	0.956	0.0041	0.0891	0.0863	0.949
$\beta=-0.5$	−0.0151	0.2056	0.2073	0.955	−0.0067	0.1361	0.1402	0.959
$\gamma=-0.5$	−0.0034	0.0854	0.0923	0.953	−0.0021	0.0612	0.0630	0.958
$\sigma^2=0.5$	−0.0019	0.1336	0.1310	0.939	−0.0087	0.0865	0.0900	0.946
$\alpha=0.5$	0.0104	0.1217	0.1258	0.952	0.0055	0.0828	0.0862	0.964
$\beta=0$	−0.0045	0.1944	0.1992	0.956	−0.0093	0.1356	0.1366	0.949
$\gamma=-0.5$	−0.0019	0.0860	0.0923	0.954	−0.0010	0.0596	0.0626	0.955
$\sigma^2=0.5$	−0.0064	0.1267	0.1314	0.945	−0.0016	0.0880	0.0913	0.959
$\alpha=0$	0.0059	0.0526	0.0565	0.972	0.0047	0.0363	0.0379	0.965
$\beta=0.5$	0.0221	0.1624	0.1670	0.960	0.0167	0.1134	0.1143	0.950
$\gamma=0.5$	0.0004	0.0950	0.0948	0.941	0.0033	0.0604	0.0642	0.957
$\sigma^2=0.5$	−0.0029	0.1368	0.1399	0.933	0.0013	0.0985	0.0968	0.943
$\alpha=0.5$	0.0113	0.1168	0.1267	0.959	0.0072	0.0856	0.0871	0.955
$\beta=0.5$	0.0130	0.2056	0.2068	0.955	−0.0073	0.1375	0.1397	0.960
$\gamma=0$	−0.0011	0.0894	0.0939	0.961	−0.0028	0.0601	0.0638	0.959
$\sigma^2=0.5$	−0.0041	0.1359	0.1361	0.943	−0.0020	0.0924	0.0940	0.941

As evidenced in Table 1, the proposed estimators appear to be unbiased, with the estimated standard errors closely aligning with the sample standard errors. Additionally, the results on the 95% coverage probabilities imply that the estimators' distribution can be

adequately approximated by a normal distribution. As expected, there is an improvement in performance with larger sample sizes, and the CPs are nearly close to the nominal level. Furthermore, Figures 1 and 2 display the estimated results of the cumulative baseline hazard functions with $\Lambda_1(t) = 0.1t, \Lambda_2(t) = t$ for sample sizes of 200 and 400, respectively. From the figures, it can be observed that for sample sizes of 200 and 400, the estimated curves of the cumulative baseline hazard functions are quite close to the true curves, indicating that the estimators are unbiased. Furthermore, as the sample size increases, the distance between the estimated curves and the true curves decreases.

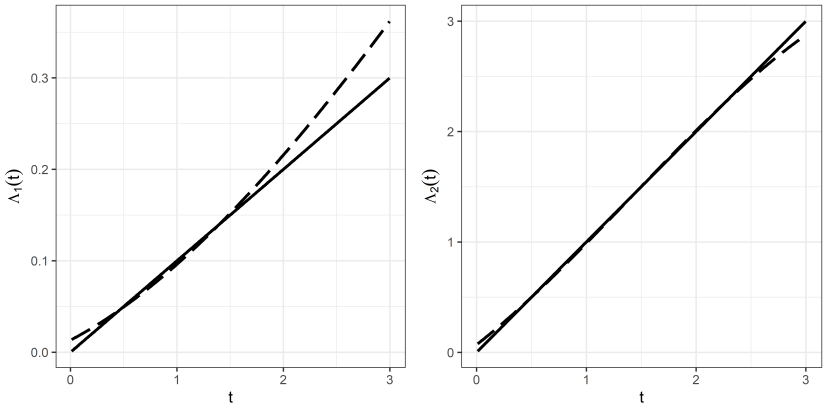

Figure 1. The simulated results of the baseline functions with a sample size of 200. (The dashed lines represent the estimated curves of the functions and the solid lines represent the true curves of the functions.).

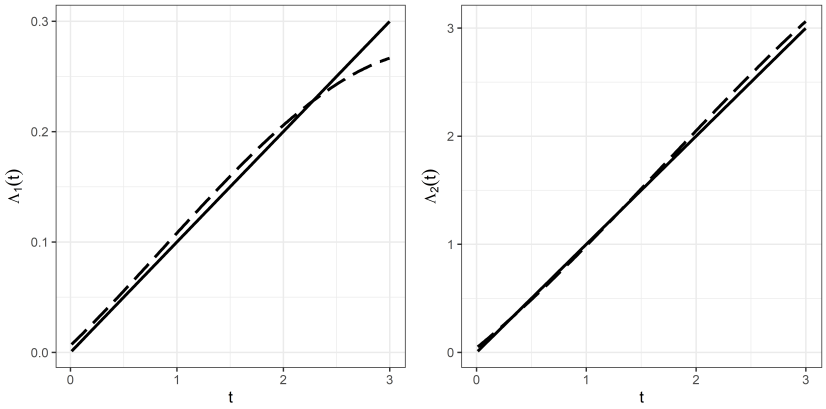

Figure 2. The simulated results of the baseline functions with a sample size of 400. (The dashed lines represent the estimated curves of the functions and the solid lines represent the true curves of the functions.).

Table 2 presents the estimation results for the two-dimensional covariate situation, and the true values of α_1 and α_2 are 0.4 or -0.4, the true values of β_1 and β_2 are 0.4, -0.4 or 0, the true values of γ are 0.4 or -0.4, and the true value of σ^2 is 0.4. The rest of the simulated settings are the same as before. The results in Tables 1 and 2 show a consistent trend, indicating that the proposed inference method also performs well in the case of two-dimensional covariates.

Table 3 presents the parameter estimation results for the case where $\sigma^2 = 0.3, 0.4, 0.5$, or 0.6, and the true values of α, β, γ are 0.5. From Table 3, it can be observed that the proposed method performs well under different variance values of the latent variables.

Table 2. Parameter estimation results under two-dimensional covariates.

Para	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
		n = 200				**n = 400**		
$\alpha_1 = 0.4$	0.0164	0.1295	0.1309	0.961	−0.0002	0.0929	0.0893	0.948
$\alpha_2 = 0.4$	0.0097	0.1307	0.1298	0.954	0.0049	0.0907	0.0897	0.956
$\beta_1 = 0$	0.0077	0.2623	0.2570	0.952	0.0021	0.1732	0.1697	0.946
$\beta_2 = 0$	0.0047	0.2666	0.2594	0.951	0.0009	0.1715	0.1703	0.946
$\gamma_1 = -0.4$	−0.0019	0.0834	0.0895	0.961	0.0010	0.0581	0.0604	0.952
$\gamma_2 = -0.4$	−0.0033	0.0851	0.0900	0.955	−0.0019	0.0631	0.0603	0.941
$\sigma^2 = 0.4$	−0.0052	0.1119	0.1150	0.936	−0.0008	0.0792	0.0787	0.940
$\alpha_1 = 0.4$	0.0136	0.1303	0.1307	0.955	0.0019	0.0934	0.0898	0.939
$\alpha_2 = 0.4$	0.0149	0.1353	0.1310	0.941	0.0056	0.0884	0.0900	0.957
$\beta_1 = -0.4$	−0.0109	0.2468	0.2614	0.963	−0.0029	0.1739	0.1709	0.942
$\beta_2 = -0.4$	−0.0064	0.2552	0.2621	0.959	−0.0031	0.1679	0.1719	0.949
$\gamma_1 = -0.4$	0.0028	0.0874	0.0899	0.956	0.0012	0.0585	0.0608	0.955
$\gamma_2 = -0.4$	0.0026	0.0868	0.0900	0.943	−0.0018	0.0631	0.0606	0.941
$\sigma^2 = 0.4$	−0.0082	0.1185	0.1144	0.931	−0.0015	0.0779	0.0786	0.944
$\alpha_1 = 0$	0.0047	0.0589	0.0549	0.948	0.0055	0.0387	0.0366	0.940
$\alpha_2 = 0$	0.0084	0.0570	0.0546	0.948	0.0055	0.0373	0.0366	0.945
$\beta_1 = 0.4$	0.0377	0.1842	0.1702	0.939	0.0214	0.1204	0.1141	0.937
$\beta_2 = 0.4$	0.0416	0.1793	0.1710	0.953	0.0282	0.1246	0.1142	0.925
$\gamma_1 = 0.4$	0.0019	0.0864	0.0910	0.956	0.0037	0.0598	0.0612	0.957
$\gamma_2 = 0.4$	−0.0024	0.0873	0.0915	0.950	0.0015	0.0626	0.0615	0.936
$\sigma^2 = 0.4$	−0.0095	0.1224	0.1244	0.929	−0.0004	0.0859	0.0853	0.941
$\alpha_1 = 0.4$	0.0122	0.1385	0.1311	0.947	0.0037	0.0914	0.0898	0.945
$\alpha_2 = 0.4$	0.0093	0.1323	0.1303	0.951	0.0051	0.0885	0.0897	0.960
$\beta_1 = 0.4$	0.0136	0.2609	0.2569	0.945	0.0056	0.1707	0.1688	0.950
$\beta_2 = 0.4$	0.0056	0.2573	0.2555	0.944	0.0023	0.1706	0.1699	0.956
$\gamma_1 = 0.4$	0.0008	0.0866	0.0909	0.959	−0.0026	0.0597	0.0612	0.957
$\gamma_2 = 0.4$	0.0016	0.0834	0.0910	0.962	0.0011	0.0582	0.0613	0.969
$\sigma^2 = 0.4$	−0.0074	0.1163	0.1145	0.930	−0.0026	0.0806	0.0783	0.936

Table 3. Parameter estimation results under different variance values of the latent variables.

Para	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
		n = 200				**n = 400**		
$\alpha = 0.5$	0.0065	0.1134	0.1166	0.955	0.0052	0.0804	0.0808	0.957
$\beta = 0.5$	0.0290	0.1888	0.1941	0.955	0.0091	0.1241	0.1317	0.953
$\gamma = 0.5$	0.0034	0.0800	0.0848	0.954	0.0014	0.0580	0.0576	0.947
$\sigma^2 = 0.3$	−0.0072	0.0999	0.1019	0.937	−0.0015	0.0690	0.0706	0.950
$\alpha = 0.5$	0.0071	0.1241	0.1218	0.945	0.0005	0.0835	0.0832	0.953
$\beta = 0.5$	0.0210	0.1970	0.2006	0.971	0.0083	0.1352	0.1356	0.948
$\gamma = 0.5$	−0.0032	0.0841	0.0897	0.957	−0.0029	0.0585	0.0607	0.958
$\sigma^2 = 0.4$	0.0013	0.1153	0.1177	0.946	−0.0054	0.0792	0.0803	0.945
$\alpha = 0.5$	0.0147	0.1274	0.1264	0.963	−0.0017	0.0849	0.0858	0.958
$\beta = 0.5$	0.0166	0.2013	0.2045	0.961	0.0078	0.1370	0.1389	0.953
$\gamma = 0.5$	0.0043	0.0867	0.0932	0.951	0.0011	0.0629	0.0638	0.955
$\sigma^2 = 0.5$	−0.0069	0.1214	0.1308	0.952	−0.0037	0.0889	0.0904	0.949
$\alpha = 0.5$	0.0101	0.1313	0.1291	0.950	0.0084	0.0874	0.0892	0.950
$\beta = 0.5$	0.0123	0.2128	0.2112	0.959	0.0115	0.1400	0.1425	0.955
$\gamma = 0.5$	0.0041	0.0924	0.0977	0.963	0.0012	0.0651	0.0664	0.941
$\sigma^2 = 0.6$	−0.0075	0.1390	0.1450	0.953	−0.0044	0.0960	0.1000	0.955

Table 4 presents the parameter estimation results for sample sizes of $n = 200$ or 400 under different cumulative baseline hazards functions, which include the cases where the cumulative hazards function is a quadratic function $\Lambda_1(t) = 0.1t^2$, $\Lambda_2(t) = 0.2t^2$, and

where the cumulative hazards function is a cubic function $\Lambda_1(t) = 0.05t^3, \Lambda_2(t) = 0.07t^3$. The other setup is the same as in Table 1. From Table 4, it can be observed that the proposed parameter estimation method performs well under different forms of cumulative hazards functions. Further investigation with different m values confirmed similar results. As evidenced in Table 5, the estimator exhibits robustness across various m selections.

Table 4. Parameter estimation results under different baseline hazards functions.

Para	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
	n = 200				**n = 400**			
			$\Lambda_1(t) = 0.1 \cdot t^2, \Lambda_2(t) = 0.2 \cdot t^2$					
$\alpha = 0.4$	−0.0078	0.0952	0.1031	0.950	−0.0052	0.0652	0.0701	0.958
$\beta = 0.4$	0.0051	0.1842	0.1911	0.955	0.0006	0.1350	0.1304	0.950
$\gamma = 0.4$	0.0018	0.0877	0.0939	0.960	0.0003	0.0611	0.0631	0.954
$\sigma^2 = 0.4$	0.0006	0.1250	0.1226	0.934	−0.0042	0.0828	0.0838	0.948
$\alpha = 0.4$	−0.0123	0.0954	0.1025	0.956	−0.0074	0.0708	0.0705	0.946
$\beta = -0.4$	0.0141	0.1872	0.1933	0.953	0.0122	0.1309	0.1313	0.947
$\gamma = -0.4$	0.0006	0.0859	0.0924	0.962	0.0036	0.0611	0.0630	0.960
$\sigma^2 = 0.4$	−0.0139	0.1154	0.1213	0.933	0.0030	0.0872	0.0847	0.937
			$\Lambda_1(t) = 0.05 \cdot t^3, \Lambda_2(t) = 0.07 \cdot t^3$					
$\alpha = 0.4$	−0.0206	0.1026	0.1050	0.944	−0.0133	0.0694	0.0715	0.949
$\beta = 0.4$	−0.0071	0.1712	0.1879	0.964	−0.0038	0.1259	0.1278	0.957
$\gamma = 0.4$	0.0068	0.0926	0.0999	0.952	0.0010	0.0667	0.0679	0.954
$\sigma^2 = 0.4$	0.0054	0.1339	0.1301	0.940	0.0052	0.0896	0.0896	0.948
$\alpha = 0.4$	−0.0229	0.0987	0.1058	0.945	−0.0183	0.0722	0.0717	0.943
$\beta = -0.4$	0.0492	0.1664	0.1884	0.957	0.0248	0.1240	0.1285	0.954
$\gamma = -0.4$	0.0123	0.0923	0.0986	0.959	0.0064	0.0654	0.0672	0.943
$\sigma^2 = 0.4$	0.0016	0.1311	0.1304	0.941	0.0039	0.0946	0.0894	0.932

Table 5. Parameter estimation results under different Bernstein polynomial degrees.

	Para	Bias	SSE	ESE	CP	Bias	SSE	ESE	CP
		n = 200				**n = 400**			
$m = 4$	$\alpha = 0.5$	0.0107	0.1306	0.1256	0.942	0.0099	0.0898	0.0872	0.945
	$\beta = 0.5$	0.0283	0.1995	0.2031	0.950	0.0106	0.1403	0.1394	0.951
	$\gamma = 0.5$	0.0030	0.0902	0.0933	0.959	0.0001	0.0608	0.0641	0.959
	$\sigma^2 = 0.5$	0.0073	0.1276	0.1307	0.943	−0.0012	0.0905	0.0914	0.959
$m = 5$	$\alpha = 0.5$	0.0081	0.1249	0.1261	0.950	0.0027	0.0845	0.0854	0.952
	$\beta = 0.5$	0.0267	0.1997	0.2052	0.954	0.0187	0.1381	0.1388	0.952
	$\gamma = 0.5$	0.0021	0.0889	0.0943	0.968	0.0004	0.0633	0.0631	0.951
	$\sigma^2 = 0.5$	−0.0021	0.1280	0.1318	0.949	−0.0007	0.0918	0.0903	0.942
$m = 6$	$\alpha = 0.5$	0.0071	0.1339	0.1272	0.944	0.0052	0.0898	0.0863	0.937
	$\beta = 0.5$	0.0072	0.2005	0.2042	0.958	0.0076	0.1409	0.1390	0.957
	$\gamma = 0.5$	0.0023	0.0887	0.0944	0.961	0.0002	0.0660	0.0637	0.945
	$\sigma^2 = 0.5$	−0.0088	0.1240	0.1321	0.953	−0.0057	0.0883	0.0900	0.949

5. An Application

In this section, we have applied the proposed procedure to analyze authentic data from a skin cancer chemoprevention study, which was undertaken by the University of Wisconsin Comprehensive Cancer Center Madison, Wisconsin, as reported in [23]. This was a five-year, double-blind, placebo-controlled, randomized Phase III clinical trial. The main goal of this trial was to assess the efficacy of a daily oral dose of 0.5 g/m^2 PO difluoromethylornithine (DFMO) in reducing the incidence of new skin cancers among patients with a past history of non-melanoma skin cancers, specifically basal cell carcinoma and squamous cell carcinoma. The participants were scheduled for biannual evaluations to

monitor the emergence of new skin cancers. The trial involved 291 participants, with 147 assigned to the placebo group and 144 to the DFMO group. The collected data encompassed the incidence rates of both basal cell carcinoma and squamous cell carcinoma at different observation points. For the timing of the first recurrence of squamous cell carcinoma and the overall recurrence pattern of basal cell carcinoma, only interval-censored data and panel count data were recorded, respectively. It should be noted that these two variables are crucial for understanding the recurrence process of carcinomas. Consequently, for a comprehensive evaluation of the treatment's impact on carcinoma recurrence, it is advisable to consider both variables simultaneously.

In this research, apart from the treatment indicator, we possess data on three additional baseline covariates for the following subjects: gender, age at diagnosis, and the count of previous skin cancers from the initial diagnosis to the time of randomization. In the following analysis, we focus on 290 patients, 147 in the placebo group and 143 in the DFMO group, who recorded at least one observation. In order to utilize the estimation procedure previously outlined, we define specific variables for each patient i (where i ranges from 1 to 290). Specifically, T_i represents the time to the first recurrence of squamous cell carcinoma, while $N_i(t)$ is the total count of basal cell carcinomas diagnosed by time t. Furthermore, we define the covariates as follows: X_{i1} denotes the patient's age, X_{i2} is an indicator variable equal to 1 if the patient is female and 0 if male, X_{i3} is an indicator variable equal to 1 if the patient is in the DFMO group and 0 if in the placebo group, and X_{i4} denotes the number of previous skin cancer occurrences.

We consider the eight possible combinations where the effects of the four covariates are either additive or multiplicative. The effect of X_{i1} is multiplicative in models AMM1, AMM3, AMM5, AMM6, and AMM8, and additive in the remaining models; the effect of X_{i2} is additive in models AMM1, AMM4, and AMM5, and multiplicative in the remaining models; the effect of X_{i3} is additive in models AMM1, AMM3, AMM5, and AMM6, and multiplicative in the remaining models; the number of previous skin cancers X_{i4} has a multiplicative effect in models AMM1, AMM6, and AMM7, and an additive effect in the remaining models. The analysis results are summarized in Table 6.

The results presented in Table 6 reveal that neither the gender at diagnosis nor the treatment method exerts a substantial effect on the time to the first squamous cell carcinoma recurrence, irrespective of whether their effects are multiplicative or additive. Moreover, there is a positive correlation between the time to the first squamous cell carcinoma recurrence and the number of prior skin cancers. However, the effect of the age at diagnosis on the time to the first squamous cell carcinoma recurrence seems to be multiplicative rather than additive and is positively correlated with the time to the first recurrence of squamous cell carcinoma. Moreover, the variances estimated for the latent variables across all the models are significantly non-zero, with p-values below the 0.05 threshold, which supports the reliability of the assumption that there is a dependency between the timing of the first squamous cell carcinoma recurrence and the recurrence process of basal cell carcinoma.

Additionally, we have also considered the use of the Cox proportional hazards model for modeling the failure time by Xu et al. [14], and analyzed the example data to obtain the coefficient estimates for covariates. The coefficient estimates for X_2 and X_3 are 0.3607 and -0.1213, respectively, with p-values of 0.12 and 0.60. The estimate for X_4 is 3.2025, with a p-value below the 0.05, which is similar to our results, whereas the coefficient estimate for X_1 is 1.5562, with a p-value of 0.01. Although the result that covariate X_1 has a significant impact on the time to the first recurrence of squamous cell carcinoma is consistent with our conclusion, our method suggests that the effect of covariate X_1 is additive rather than multiplicative by comparing the estimated results across different models.

Table 6. Parameter estimation results of the skin cancer data.

	Gender			**Group**		
	EST	**ESE**	***p*-Value**	**EST**	**ESE**	***p*-Value**
AMM1	0.0533	0.0407	0.1903	−0.0207	0.0419	0.6212
AMM2	0.2919	0.3617	0.4196	−0.1098	0.2656	0.6793
AMM3	−0.4050	0.3859	0.2940	−0.0031	0.0380	0.9348
AMM4	0.0544	0.0441	0.2179	−1.5994	4.3731	0.7146
AMM5	0.0532	0.0431	0.2167	−0.0310	0.0432	0.4719
AMM6	0.4467	0.2294	0.0515	−0.0096	0.0380	0.8007
AMM7	0.2406	0.1922	0.2108	0.0130	0.1656	0.9374
AMM8	0.4194	0.5531	0.4483	−0.4620	0.4964	0.3519
	age			the number of prior cancers		
	EST	ESE	*p*-value	EST	ESE	*p*-value
AMM1	0.4386	0.1639	0.0075	0.5612	0.1147	<0.0001
AMM2	−0.0308	0.1390	0.8248	0.4848	0.2796	0.0830
AMM3	0.8935	0.3721	0.0163	0.7565	0.2698	0.0050
AMM4	0.0153	0.0748	0.8384	1.0160	0.3325	0.0022
AMM5	0.7329	0.4469	0.1010	0.7461	0.3033	0.0139
AMM6	0.4398	0.1479	0.0029	0.5243	0.0955	<0.0001
AMM7	−0.0558	0.1178	0.6355	0.3722	0.1461	0.0108
AMM8	1.4338	0.5443	0.0084	1.0880	0.2611	<0.0001

Figures 3 and 4 present the estimation results of the baseline functions for all the models, respectively. From Figure 4, it is evident that the estimated curves for $\Lambda_2(t)$ across all the models are highly congruent. This congruence can be attributed to the similarity in the estimation outcomes of γ across the models. In contrast, for $\Lambda_1(t)$, there is a divergence in the estimation curves among the different models. This variation may be attributed to the distinct mechanisms by which covariates influence the failure time across models, which can be either additive or multiplicative. Moreover, the significance of the impact of the same covariate across various models also exhibits notable differences.

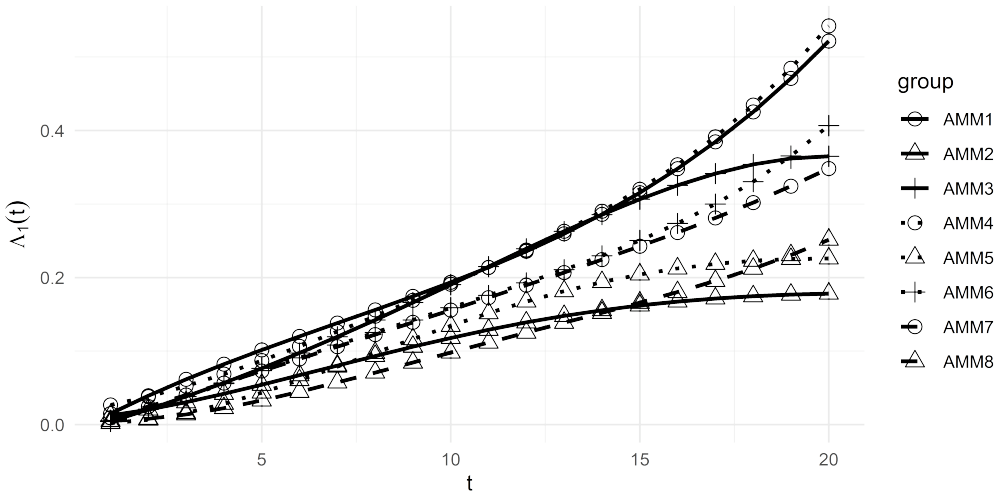

Figure 3. The results of the baseline hazard function $\Lambda_1(t)$.

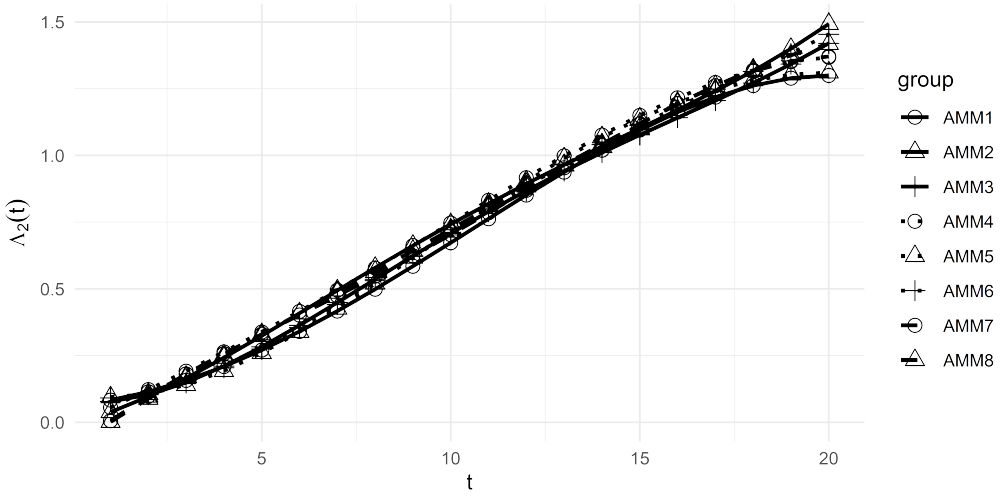

Figure 4. The results of the baseline mean function $\Lambda_2(t)$.

6. Discussion

The present study explores the joint analysis of interval-censored data and panel count data. It assumes that the hazards function of the failure time follows an additive–multiplicative model. A joint model incorporating frailty is suggested to model the failure time of interest and the panel count process. For the nonparametric part of the model, Bernstein polynomials are used for approximation, and the final parameter estimation is obtained through the method of maximum likelihood estimation. Finally, simulation studies and case analyses are conducted. The results of the simulation and case studies demonstrate that the method is effective in practical situations, and the proposed estimators are consistent and asymptotically normal.

In addition, the method proposed relies on certain assumptions and has some limitations. The first hypothesis is that the process of the recurrent event follows a Poisson distribution, a theory that Wellner and Zhang [5] have confirmed to offer robustness in the estimation process under comparable scenarios. Although the nonhomogeneous Poisson process is widely applied in many fields, including reliability engineering, ecology, and epidemiology, due to its flexibility and adaptability to changes over time, and it can simulate changes in risk or event rates across different time periods, in some cases, actual data may exhibit greater variability than what is predicted by the nonhomogeneous Poisson process. And we also assume that the mean function of the count process follows a proportional mean model; it is meaningful to extend this assumption to other more general models.

Another assumption of the method proposed is that the observation process is non-informative. However, in many practical problems, this assumption may be violated. Consequently, further research should consider incorporating useful information from the observation process into modeling, while jointly modeling the failure time process, the counting process of recurrent events, and the observation process. This study assumes that the frailties follow a gamma distribution, which simplifies the form of the likelihood function. When the frailties follow other general distributions, the computation of the expectation with respect to the frailties is more complex, but numerical computation methods such as Gaussian quadrature or Monte Carlo integration can be considered, thus allowing the proposed method to be generalized to the general case where frailties follow other distributions. In situations where there is no distributional assumption for the frailties, new modeling and corresponding estimation methods need to be developed, such as the two-step method; refer to Wang et al. [24] and Wang et al. [25].

Noting the involvement of squamous cell carcinoma and basal cell carcinoma in the analysis of skin cancer data, our study focuses on the failure time of the event of interest, specifically the first occurrence of squamous cell carcinoma, while also considering the impact of the overall recurrence pattern of basal cell carcinoma. Further research considering the joint analysis of the first occurrence of basal cell carcinoma and the recurrence of squamous cell carcinoma is equally significant.

Author Contributions: Writing—original draft preparation, T.W.; methodology, J.S.; software, T.W.; formal analysis, Y.L.; modeling, S.W.; writing—review and editing, J.S. and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by the Outstanding Youth Foundation of Jilin Natural Science Foundation (20230101371JC).

Data Availability Statement: The original data used in the empirical study is available in Table A.3 of Sun and Zhao [23].

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

References

1. Zhu, L.; Tong, X.; Cai, D.; Li, Y.; Sun, R.; Srivastava, D.K.; Hudson, M.M. Maximum likelihood estimation for the proportional odds model with mixed interval-censored failure time data. *J. Appl. Stat.* **2021**, *48*, 1496–1512. [CrossRef] [PubMed]
2. Sun, T.; Ding, Y. Copula-based semiparametric regression method for bivariate data under general interval censoring. *Biostatistics* **2021**, *22*, 315–330. [CrossRef] [PubMed]
3. Bouaziz, O.; Lauridsen, E.; Nuel, G. Regression modelling of interval censored data based on the adaptive ridge procedure. *J. Appl. Stat.* **2022**, *49*, 3319–3343. [CrossRef] [PubMed]
4. Sun, J.; Kalbfleisch, J.D. Estimation of the mean function of point processes based on panel count data. *Stat. Sin.* **1995**, *1995*, 279–289.
5. Wellner, J.A.; Zhang, Y. Two estimators of the mean of a counting process with panel count data. *Ann. Stat.* **2000**, *28*, 779–814. [CrossRef]
6. Hu, X.J.; Lagakos, S.W.; Lockhart, R.A. Marginal analysis of panel counts through estimating functions. *Biometrika* **2009**, *96*, 445–456. [CrossRef]
7. Hua, L.; Zhang, Y. Spline-based semiparametric projected generalized estimating equation method for panel count data. *Biostatistics* **2012**, *13*, 440–454. [CrossRef]
8. Hua, L.; Zhang, Y.; Tu, W. A spline-based semiparametric sieve likelihood method for over-dispersed panel count data. *Can. J. Stat.* **2014**, *42*, 217–245. [CrossRef]
9. Yao, B.; Wang, L.; He, X. Semiparametric regression analysis of panel count data allowing for within-subject correlation. *Comput. Stat. Data Anal.* **2016**, *97*, 47–59. [CrossRef]
10. Wellner, J.A.; Zhang, Y. Two likelihood-based semiparametric estimation methods for panel count data with covariates. *Ann. Stat.* **2007**, *35*, 2106–2142. [CrossRef]
11. Wen, C.C.; Chen, Y.H. Joint analysis of current count and current status data. *J. Multivar. Anal.* **2016**, *143*, 153–164. [CrossRef]
12. Wen, C.C.; Chen, Y.H. Pseudo and conditional score approach to joint analysis of current count and current status data. *Biometrics* **2018**, *74*, 1223–1231. [CrossRef] [PubMed]
13. Wen, C.C.; Chen, Y.H.; Tseng, C.H. Joint analysis of panel count and interval-censored data using distribution-free frailty analysis. *Biom. J.* **2020**, *62*, 1164–1175. [CrossRef] [PubMed]
14. Xu, D.; Zhao, H.; Sun, J. Joint analysis of interval-censored failure time data and panel count data. *Lifetime Data Anal.* **2018**, *24*, 94–109. [CrossRef]
15. Lin, D.Y.; Ying, Z. Semiparametric analysis of general additive-multiplicative hazard models for counting processes. *Ann. Stat.* **1995**, *1995*, 1712–1734. [CrossRef]
16. Chen, X. Large sample sieve estimation of semi-nonparametric models. *Handb. Econom.* **2007**, *6*, 5549–5632.
17. Carnicer, J.M.; Pena, J.M. Shape preserving representations and optimality of the Bernstein basis. *Adv. Comput. Math.* **1993**, *1*, 173–196. [CrossRef]
18. Huang, J.; Rossini, A.J. Sieve estimation for the proportional-odds failure-time regression model with interval censoring. *J. Am. Stat. Assoc.* **1997**, *92*, 960–967. [CrossRef]
19. Lorentz, G. *Bernstein Polynomials*, 2nd ed.; University Toronto Press: Toronto, ON, Canada, 1986.
20. Shen, X. On methods of sieves and penalization. *Ann. Stat.* **1997**, *25*, 2555–2591. [CrossRef]
21. Murphy, S.A.; Van der Vaart, A.W. On profile likelihood. *J. Am. Stat. Assoc.* **2000**, *95*, 449–465. [CrossRef]
22. Zeng, D.; Cai, J.; Shen, Y. Semiparametric additive risks model for interval-censored data. *Stat. Sin.* **2006**, *16*, 287–302.
23. Sun, J.; Zhao, X. *Statistical Analysis of Panel Count Data*; Springer: New York, NY, USA, 2013.

24. Wang, P.; Zhao, H.; Sun, J. Regression analysis of case K interval-censored failure time data in the presence of informative censoring. *Biometrics* **2016**, *72*, 1103–1112. [CrossRef] [PubMed]
25. Wang, S.; Wang, C.; Song, X.; Xu, D. Joint analysis of informatively interval-censored failure time and panel count data. *Stat. Methods Med. Res.* **2022**, *31*, 2054–2068. [CrossRef] [PubMed]