Biometrics 000, 000–000 DOI: 000
000 0000

# Semiparametric Estimation of the Accelerated Mean Model with Panel Count Data under Informative Examination Times

**Sy Han Chiou[1,*], Gongjun Xu[2], Jun Yan[3], and Chiung-Yu Huang[4]**

[1] Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75080, U.S.A.

[2] Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, U.S.A.

[3] Department of Statistics, University of Connecticut, Storrs, Connecticut 06269, U.S.A.

[4] Department of Epidemiology and Biostatistics, University of California San Francisco
San Francisco, California 94158, U.S.A.

*email:* schiou@utdallas.edu



Summary: Panel count data arise when the number of recurrent events experienced by each subject is observed intermittently at discrete examination times. The examination time process can be informative about the underlying recurrent event process even after conditioning on covariates. We consider a semiparametric accelerated mean model for the recurrent event process and allow the two processes to be correlated through a shared frailty. The regression parameters have a simple marginal interpretation of modifying the time scale of the cumulative mean function of the event process. A novel estimation procedure for the regression parameters and the baseline rate function is proposed based on a conditioning technique. In contrast to existing methods, the proposed method is robust in the sense that it requires neither the strong Poisson-type assumption for the underlying recurrent event process nor a parametric assumption on the distribution of the unobserved frailty. Moreover, the distribution of the examination time process is left unspecified, allowing for arbitrary dependence between the two processes. Asymptotic consistency of the estimator is established, and the variance of the estimator is estimated by a model-based smoothed bootstrap procedure. Numerical studies demonstrated that the proposed point estimator and variance estimator perform well with practical sample sizes. The methods are applied to data from a skin cancer chemoprevention trial.

Key words: Frailty; Model-based bootstrap; Poisson process; Recurrent events; Scale-change model; Squared extrapolation method.

## 1. Introduction

Panel count data arise when recurrent events are examined periodically rather than continuously due to cost, feasibility, or other practical considerations (Kalbfleisch and Lawless, 1985; Thall and Lachin, 1988); see Sun and Zhao (2013) for a recent review. As a result, instead of the exact event times, only the numbers of events that occur between successive examination times are observed. In most applications, the examination times may depend on the underlying risk of recurrent events, leading to so-called informative examination times. For example, in a skin cancer chemoprevention clinical trial, many patients have multiple recurrences of skin tumors throughout the study, but occurrences of new tumors were observed only at clinical visits (Bailey et al., 2010). Exploratory data analyses suggested that patients with higher tumor recurrence rates tend to have more frequent clinical visits as they may require more medical attention (Li et al., 2011; Sun and Zhao, 2013). In another example, Ma and Sundaram (2016) studied the labor progression of women who had no previous birth experience by treating each 1 cm increment of cervical dilation as a recurrent event. During labor, vaginal examinations are performed at intermittent time points to assess for cervical dilation, so only event counts are observed. Obviously, the timing and frequency of examination are correlated with the dilation process; the faster the cervix dilates, the more frequently a woman is getting examined. Negative dependence between the recurrent event process and the examination process may be possible in other applications. As pointed out by many authors (e.g., Huang et al., 2006; Sun et al., 2007), statistical methods that fail to account for such dependency can yield substantial bias and misleading inferential results.

When covariate effects are of interest, Cox-type models are commonly used. The majority of the earlier literature on panel count data analysis assumed uninformative examination times, that is, the examination time process is independent of the recurrent event process given covariates. For example, Zhang (2002), Wellner and Zhang (2007), and Lu et al. (2009)

considered pseudo-likelihood and likelihood methods under the nonhomogenous Poisson process assumption. They showed that both methods are robust against departure from the Poisson assumption as long as the proportional rates model holds. Sun and Wei (2000) and Hu et al. (2003) considered estimating equation approaches based on cumulative event counts at different time points. The estimating equation approaches are computationally more convenient but can be inefficient; improvement in efficiency is possible in certain situations through generalized estimating equations (Hua and Zhang, 2012).

The need to develop statistical methods that can deal with informative examination times has been increasing recognized. Kim (2006) fully specified both the recurrent event process and the examination time processes with a shared gamma frailty. Authors, including Sun et al. (2007), He et al. (2009), Zhao and Tong (2011), and Zhao et al. (2013), have extended the methodology proposed in Sun and Wei (2000) to allow the two processes to be correlated through a shared frailty with an unspecified distribution. Extending the estimation equation-based method of Zeng and Cai (2010), Zhou et al. (2017) considered a flexible joint model of the recurrent event process, the examination time process, and the time to a terminal event, where the associations between processes are left unspecified. Buzkova (2010) proposed to model the dependency of examination time process on the history of observed recurrent event counts, thus permits outcome-dependent examination times; the inverse-intensity-rate-ratio weighting technique (Buzkova and Lumley, 2007) was applied to construct unbiased estimating equations. Naturally, the validity of the aforementioned methods rely on correct model specifications for the examination time process and the follow-up time (or a terminal event time), which may not be of primary interest in practice. In contrast, Huang et al. (2006) and Wang et al. (2013) postulated a frailty proportional rates model for the recurrent event process, where the distributions of the frailty and the possibly correlated examination times are left unspecified. Their estimation procedures eliminate the nuisance frailties through a

conditioning technique and the resulting estimators are robust against departure from the Poisson assumption on the event processes.

As an alternative to the Cox-type formulation, we propose an accelerated mean model for the recurrent event process under informative examination times. This is a new framework compared to other attempts to go beyond the Cox-type formulation such as the semiparametric transformation models studied in Li et al. (2010) and Li et al. (2013), where a correct model specification for the dependency of cumulative event count on the history of the examination times is required, and, more importantly, the regression parameters of the covariates of interest can be less intuitive to interpret. Motivated by the accelerated failure time (AFT) model for recurrent event processes (e.g., Lin et al., 1998; Xu et al., 2017), we assume that the covariates have a time-scale-change effect on the marginal mean cumulative function. The examination process is allowed to be informative about the recurrent event process through a subject-specific multiplicative frailty. The distribution of the frailty is left unspecified because our estimation procedure eliminates the unobserved frailty via a conditioning approach in a way similar to that of Wang et al. (2001) and Huang et al. (2006). Unconditional on the frailty, the model allows for an unspecified association between the recurrent event process and the examination time process. No model is needed for the examination time process, an appealing feature when it is not of primary interest.

We proposed a novel estimation procedure that iterates between updating the cumulative baseline rate function and updating the regression parameter. The squared extrapolation method (SQUAREM) of Varadhan and Roland (2008) is adopted to accelerate the expectation-maximization type algorithm in estimating the cumulative baseline rate function in each iteration. To our knowledge, this is the first time it is applied to semiparametric estimation; in our case, it increased the speed by a factor of 5 on average. The consistency of the estimator is established under suitable regularity conditions without the Poisson

assumption on the recurrent event process. For variance estimation, we propose a model-based smoothed bootstrap procedure motivated by Sen and Xu (2015) to provide better coverage probabilities than the standard nonparametric bootstrap procedure. The methods are applied to the skin cancer example along with a goodness of fit assessment.

## 2. Semiparametric Accelerated Mean Model

### 2.1 *Model Setup*

Consider panel count data observed in a fixed time interval $[0, \tau]$ from $n$ independent subjects. For the $i$th subject, let $N_i(t)$ be the number of events over the interval $[0, t]$, and $\boldsymbol{X}_i$ be a $p \times 1$ covariate vector. We assume the event process $N_i(\cdot)$ of the $i$th subject is only observable at $K_i$ discrete random time points, $0 = t_{i0} < t_{i1} < t_{i2} < \ldots < t_{iK_i} \leqslant \tau$, where $t_{ij}$ is the $j$th examination time, $j = 1, \ldots, K_i$. Suppose that the last examination time $t_{iK_i}$ is also the follow-up time of subject $i$. The observed panel count data are a random sample $\{t_{ij}, K_i, N_i(t_{ij}), \boldsymbol{X}_i; j = 1, \ldots, K_i\}$, $i = 1, \ldots, n$.

As in Xu et al. (2017), we assume that the recurrent event process $N_i(\cdot)$, conditioning on a latent nonnegative frailty variable $Z_i$ and covariate $\boldsymbol{X}_i$, has the rate function

$$\lambda_i(t) = Z_i \lambda_0(te^{\boldsymbol{X}_i^\top \boldsymbol{\alpha}}) e^{\boldsymbol{X}_i^\top \boldsymbol{\alpha}}, \quad t \in [0, \tau], \tag{1}$$

where $\boldsymbol{\alpha}$ is a $p \times 1$ vector of parameters and $\lambda_0(t)$ is an unspecified, absolutely continuous baseline rate function. Given $Z_i$ and $\boldsymbol{X}_i$, the event process $N_i(\cdot)$ is assumed to be independent of the number of examination time points $K_i$, and the examination times $\{t_{i1}, \ldots, t_{iK_i}\}$. This allows $N_i(\cdot)$ to be dependent on $\{t_{i1}, \ldots, t_{iK_i}\}$ through unobserved frailty $Z_i$ after conditioning on $\boldsymbol{X}_i$. From Model (1), one can derive the conditional mean:

$$E\{N_i(t) \mid \boldsymbol{X}_i, Z_i\} = Z_i \Lambda_0(te^{\boldsymbol{X}_i^\top \boldsymbol{\alpha}}), \tag{2}$$

where $\Lambda_0(t) = \int_0^t \lambda_0(u)\, \mathrm{d}u$. The effect of the covariates is a scale change on the time of the

cumulative mean function of the underlying event process, which is why the model is referred to as an *accelerated mean model*.

In contrast to most joint modeling approaches (e.g., He et al., 2009), no Poisson-type assumption is imposed on $N_i(\cdot)$. Moreover, both the distribution of $Z_i$'s and the conditional distribution of the examination times given $Z_i$ are left unspecified. For model identifiability, we assume $E(Z_i|\boldsymbol{X}_i) = 1$. Then, unconditional on $Z_i$, the cumulative mean function of $N_i(\cdot)$ is $E\{N_i(t)|\boldsymbol{X}_i\} = \Lambda_0(te^{\boldsymbol{X}_i^\top \boldsymbol{\alpha}})$, which is also of the form of an accelerated mean model. In a two-arm clinical trial, for example, $\boldsymbol{\alpha}$ identifies the time scale change of the cumulative mean function in the treated group ($\boldsymbol{X}_i = 1$); the expected number of events by time $t$ among treated subjects equals the expected number of events by time $te^{\alpha}$ in the control group ($\boldsymbol{X}_i = 0$), with other risk factors being held the same. The accelerated mean model is an extension of the AFT model in the recurrent event setting. Let $U_{ij}$ be the time of the $j$th recurrent event from subject $i$, it can be shown that $\log U_{ij} = -\boldsymbol{X}_i^\top \boldsymbol{\alpha} + \epsilon_{ij}$, where the independent error vectors $(\epsilon_{ij} : j = 1, 2, \ldots)$, $i = 1, \ldots, n$, follow a common unspecified joint distribution (Lin et al., 1998; Ghosh and Lin, 2003).

## 2.2 *Point Estimation*

We first consider point estimation for the regression parameter $\boldsymbol{\alpha}$. For any $p \times 1$ vector $\boldsymbol{a}$, consider the transformation $t_{ij}^*(\boldsymbol{a}) = t_{ij}e^{\boldsymbol{X}_i^\top \boldsymbol{a}}$, $i = 1, \ldots, n$, $j = 1, \ldots, K_i$. Let $Y_i = t_{iK_i}$ and $Y_i^*(\boldsymbol{a}) = Y_i e^{\boldsymbol{X}_i^\top \boldsymbol{a}}$. Suppose $\boldsymbol{X}_i$ is bounded as in Condition C2 of the Supplementary Materials, define $\tau_{n,\boldsymbol{a}} = \tau \sup_i e^{\boldsymbol{X}_i^\top \boldsymbol{a}}$ and assume $\tau_{n,\boldsymbol{a}} \to \tau_{\boldsymbol{a}} < \infty$ as $n \to \infty$. Let $N_i^*(t, \boldsymbol{a})$ be the counting process on the transformed time scale corresponding to the original underlying event process $N_i(t)$. Then, unconditional on $Z_i$ the cumulative rate function of $N_i^*(t, \boldsymbol{a})$ is

$$E\{N_i^*(t, \boldsymbol{a}) \mid \boldsymbol{X}_i\} = E\left[E\{N_i^*(t, \boldsymbol{a}) \mid Z_i, \boldsymbol{X}_i\} \mid \boldsymbol{X}_i\right] = \Lambda_0\{te^{\boldsymbol{X}_i^\top (\boldsymbol{\alpha} - \boldsymbol{a})}\}, \quad t \in [0, \tau e^{\boldsymbol{X}_i^\top \boldsymbol{a}}]. \quad (3)$$

We use the property that the cumulative rate function of $N_i^*(t, \boldsymbol{a})$ does not depend on $\boldsymbol{X}_i$ when $\boldsymbol{a} = \boldsymbol{\alpha}$ to construct a robust estimation procedure for $\Lambda_0(\cdot)$.

For subject $i$, let $m_{ij} = N_i(t_{ij}) - N_i(t_{ij-1})$ be the number of events in the time interval $(t_{ij-1}, t_{ij}]$ and $m_i = N_i(Y_i)$ be the total number of observed events. To better illustrate our idea, consider a working model for the moment where, conditioning on $Z_i$ and $\boldsymbol{X}_i$, the event process $N_i(\cdot)$ is a Poisson process with intensity (1). Then $m_{ij}$ is a Poisson random variable with mean $\int_{t_{ij-1}}^{t_{ij}} \lambda_i(u)\,\mathrm{d}u = Z_i\Lambda_0\{t^*_{ij}(\boldsymbol{\alpha})\} - Z_i\Lambda_0\{t^*_{ij-1}(\boldsymbol{\alpha})\}$. Conditioning on $Z_i$, $\boldsymbol{X}_i$, $m_i$, and the $K_i$ examination times, the conditional likelihood based on the observed event count data is

$$L_c(\Phi, \boldsymbol{\alpha}) \propto \prod_{i=1}^{n}\prod_{j=1}^{K_i}\left[\frac{Z_i\Lambda_0\{t^*_{ij}(\boldsymbol{\alpha})\} - Z_i\Lambda_0\{t^*_{ij-1}(\boldsymbol{\alpha})\}}{Z_i\Lambda_0\{Y^*_i(\boldsymbol{\alpha})\}}\right]^{m_{ij}} = \prod_{i=1}^{n}\prod_{j=1}^{K_i}\left[\frac{\Phi\{t^*_{ij}(\boldsymbol{\alpha})\} - \Phi\{t^*_{ij-1}(\boldsymbol{\alpha})\}}{\Phi\{Y^*_i(\boldsymbol{\alpha})\}}\right]^{m_{ij}},$$

where $\Phi(t) = \Lambda_0(t)/\Lambda_0(\tau_{\boldsymbol{\alpha}})$ defines a proper distribution function on $t \in [0, \tau_{\boldsymbol{\alpha}}]$. The conditional working likelihood, $L_c(\Phi, \boldsymbol{\alpha})$, eliminates the frailty variable $Z_i$ and is equivalent to the likelihood of a set of independently interval-censored and right-truncated data. To see this, consider a hypothetical set of independent random variables $\{U_{ijk}, i = 1, \ldots, n, j = 1, \ldots, K_i, k = 1, \ldots, m_{ij}\}$ whose distribution function is $\Phi(t)$. Assume that $U_{ijk}$ is independently right truncated by $Y^*_i(\boldsymbol{\alpha})$ and interval censored in $(t^*_{ij-1}(\boldsymbol{\alpha}), t^*_{ij}(\boldsymbol{\alpha})]$. Then its contribution to the likelihood function is $[\Phi\{t^*_{ij}(\boldsymbol{\alpha})\} - \Phi\{t^*_{ij-1}(\boldsymbol{\alpha})\}]/\Phi\{Y^*_i(\boldsymbol{\alpha})\}$, and the likelihood of the hypothetical data coincides with $L_c(\Phi, \boldsymbol{\alpha})$. Thus, given $\boldsymbol{\alpha}$, the working nonparametric maximum (conditional) likelihood estimator (NPMLE) of the distribution function $\Phi(\cdot)$ can be obtained by maximizing the conditional likelihood $L_c(\Phi, \boldsymbol{\alpha})$, which motivates the self-consistent algorithm (Turnbull, 1976) described below.

Given $\boldsymbol{\alpha}$, define the working NPMLE of $\Phi(\cdot)$ by $\widehat{\Phi}_n(\boldsymbol{\alpha}, \cdot)$. Let $0 = t_{(0)} < t_{(1)} < \ldots < t_{(L)} \leqslant \tau_{\boldsymbol{\alpha}}$ be the ordered, distinct values of the observed examination times $\{t^*_{ij}(\boldsymbol{\alpha}); K_i > 1, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant K_i\}$. For $k = 1, \ldots, L$, define $a_{ijk} = I\{t_{(k-1)} \leqslant t^*_{ij-1}(\boldsymbol{\alpha}),\ t^*_{ij}(\boldsymbol{\alpha}) \leqslant t_{(k)}\}$ and $b_{ik} = I\{t_{(k)}(\boldsymbol{\alpha}) \leqslant Y^*_i(\boldsymbol{\alpha})\}$, where $I(\cdot)$ is the indicator function. Given the estimate $\widehat{\Phi}^{(l)}_n(\boldsymbol{\alpha}, \cdot)$ at the $l$th iteration, the updated estimate is obtained by $\widehat{\Phi}^{(l+1)}_n(\boldsymbol{\alpha}, t) = \sum_{k:t_{(k)}\leqslant t} d^{(l)}_k \,/\, \sum_{k=1}^{L} d^{(l)}_k$,

where

$$d_k^{(l)} = \sum_{i=1}^{n}\sum_{j=1}^{k_i} m_{ij}\left\{\frac{a_{ijk}p_k^{(l)}}{\sum_{k=1}^{L} a_{ijk}p_k^{(l)}} + \frac{(1-b_{ik})p_k^{(l)}}{\sum_{k=1}^{L} b_{ik}p_k^{(l)}}\right\},$$

and $p_k^{(l)} = \widehat{\Phi}_n^{(l)}(\boldsymbol{\alpha}, t_{(k)}) - \widehat{\Phi}_n^{(l)}(\boldsymbol{\alpha}, t_{(k-1)})$. At convergence, $\Lambda_0(\tau_{\boldsymbol{\alpha}})$ can be estimated by $\widehat{\Lambda}_n(\boldsymbol{\alpha}, \tau_{\boldsymbol{\alpha}}) = n^{-1}\sum_{i=1}^{n} m_i / \widehat{\Phi}_n\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\}$, because Equation (3) implies that

$$\begin{aligned} E\left[m_i\Phi\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\}^{-1} \mid \boldsymbol{X}_i\right] &= E[E\{m_i\Phi\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\}^{-1} \mid Y_i, Z_i, \boldsymbol{X}_i\} \mid \boldsymbol{X}_i] \\ &= E[\Lambda_0\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\}\Phi\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\}^{-1} \mid \boldsymbol{X}_i] = \Lambda_0(\tau_{\boldsymbol{\alpha}}). \end{aligned} \tag{4}$$

This further implies that $\Lambda_0(t)$ can be estimated by $\widehat{\Lambda}_n(\boldsymbol{\alpha}, t) = \widehat{\Phi}_n(\boldsymbol{\alpha}, t)\widehat{\Lambda}_n(\tau_{\boldsymbol{\alpha}})$ from the relationship $\phi(t) = \lambda_0(t)/\Lambda_0(\tau_{\boldsymbol{\alpha}})$. Since the conditional likelihood function, $L_c$, is free from $Z_i$, the estimation of $\Phi(\cdot)$ does not require information from $Z_i$. Even though the above estimation method is constructed based on the working Poisson assumption, we show in Theorem 1 that $\widehat{\Lambda}_n(\boldsymbol{\alpha}, t)$ is consistent even without the Poisson assumption.

We now consider the estimation of the parameter $\boldsymbol{\alpha}$. It follows from Equations (3) and (4) that when $\boldsymbol{a} = \boldsymbol{\alpha}$,

$$E\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i\left[m_i\Phi^{-1}\{Y_i^*(\boldsymbol{\alpha})\} - \Lambda_0(\tau_{\boldsymbol{\alpha}})\right]\right) = 0.$$

The estimator of $\Lambda_0(\tau_{\boldsymbol{\alpha}})$ suggests an estimating equation for $\boldsymbol{\alpha}$:

$$S_n(\boldsymbol{a}) = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i\left[m_i\widehat{\Phi}_n^{-1}\{\boldsymbol{a}, Y_i^*(\boldsymbol{a})\} - \frac{1}{n}\sum_{j=1}^{n} m_j\widehat{\Phi}_n^{-1}\{\boldsymbol{a}, Y_j^*(\boldsymbol{a})\}\right] = 0. \tag{5}$$

The solution to (5), denoted by $\widehat{\boldsymbol{\alpha}}_n$, is our estimator of $\boldsymbol{\alpha}$.

To solve (5), we use a derivative-free Barzilai–Borwein spectral method (Barzilai and Borwein, 1988; La Cruz et al., 2006) that updates the estimate at iteration $s$ by an increment of the form $\gamma_n^{(s)}\delta_n^{(s)}$, where $\gamma_n^{(s)}$ is a scalar spectral steplength and $\delta_n^{(s)}$ is a line search direction.

The estimation algorithm for $\widehat{\boldsymbol{\alpha}}$ is summarized below:

Step 1 Set the initial value for $\boldsymbol{\alpha}$ by $\widehat{\boldsymbol{\alpha}}_n^{(0)}$ and $\Phi(\cdot)$ by $\widehat{\Phi}_n^{(0)}(\widehat{\boldsymbol{\alpha}}_n^{(0)}, t_{(k)}) = k/L$.

Step 2 Repeat $\widehat{\Phi}_n^{(l+1)}(\widehat{\boldsymbol{\alpha}}_n^{(l)}, t) = \sum_{k:t_{(k)}\leqslant t} d_k^{(l)} / \sum_{k=1}^{L} d_k^{(l)}$ until convergence.

Step 3 Update $\widehat{\boldsymbol{\alpha}}_n^{(s+1)} = \widehat{\boldsymbol{\alpha}}_n^{(s)} + \gamma_n^{(s)}\delta_n^{(s)}$, where $\gamma_n^{(s)}$ is the steplength and $\delta_n^{(s)}$ is the search direction.

Step 4 Repeat Step Step 2 and Step 3 until convergence.

The initial value $\widehat{\boldsymbol{\alpha}}_n^{(0)}$ can be set to zero or random. We fixed $\widehat{\boldsymbol{\alpha}}_n^{(0)} = 0$ in our implementation as our exploration results in negligible differences. We used the SQUAREM implemented in Varadhan (2014) to accelerate the repetitive estimation of $\widehat{\Phi}_n(\cdot)$ at each update of $\widehat{\boldsymbol{\alpha}}_n^{(s)}$ (Step 2). Upon successful convergence, we used the derivative-free Barzilai–Borwein spectral algorithm of Varadhan and Gilbert (2009) to update $\widehat{\boldsymbol{\alpha}}_n^{(s)}$ (Step 3). The convergence criterion was based on the $\ell$-2 norm with a prefixed tolerance of 0.001 in both Step 2 and Step 4.

The estimation procedure worked fine most of the times in our simulation study, but numerical issues arose occasionally. This is likely to be caused by the existence of very short follow-up time on the transformed scale $Y_i^*$ and nonzero $m_i$, in which case $m_i/\widehat{\Phi}_n\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\}$ would explode. We consider a heuristic adjustment that replaces $m_i/\widehat{\Phi}_n\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\}$ with $(m_i + 0.01)/[\widehat{\Phi}_n\{\boldsymbol{\alpha}, Y_i^*(\boldsymbol{\alpha})\} + 0.01]$ in Equation (5) as suggested by Wang et al. (2013). With the adjustments, the portion of non-converged replicates was less than 5% in smaller sample size scenarios ($n = 50$); the convergence was less of an issue for larger sample size ($n = 100$).

### 2.3 *Consistency Results and Resampling Methods for Inference*

We have the following consistency result for $\widehat{\boldsymbol{\alpha}}_n$ and $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, \cdot)$ with proof and necessary regular condition provide in the Supplementary Materials.

THEOREM 1: *Given conditions C1–C4 and distance $d$ between two functions defined in the Supplementary Materials, $\widehat{\boldsymbol{\alpha}}_n \to \boldsymbol{\alpha}$ and $d\{\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, t)1(t \in [0, c]), \Lambda_0(t)1(t \in [0, c])\} \to 0$, for any $c < \tau_{\boldsymbol{\alpha}}$, almost surely as $n \to \infty$.*

The convergences of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, \cdot)$ does not achieve the standard $n^{1/2}$-convergence rate, and the asymptotic distribution of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, \cdot)$ does not follow the usual Gaussian type distributions.

To illustrate the idea, first consider the ideal case when $\boldsymbol{\alpha}$ is known. As in Section 2.2, $\widehat{\Phi}_n(\boldsymbol{\alpha}, \cdot)$ is based on interval censored data with examination times $\{t^*_{ij}(\boldsymbol{\alpha}); K_i > 1, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant K_i\}$. In general, $\widehat{\Phi}_n(\boldsymbol{\alpha}, t)$ at a fixed time $t$ does not have $n^{1/2}$-convergence rate. For instance, in the current status data with $K_i = 1$, $n^{1/3}\{\widehat{\Phi}(\boldsymbol{\alpha}, t) - \Phi(t)\} \xrightarrow{d} \kappa\mathbb{C}$, where $\kappa$ is some constant depending on the derivative function of $\Phi(t)$, $\mathbb{C} = \arg\min_h\{\mathbb{Z}(h) + h^2\}$, and $\mathbb{Z}$ is a standard two-sided Brownian motion process, originating from 0. In the general mixed case interval censoring setting, the limiting distribution of $\widehat{\Phi}_n(\boldsymbol{\alpha}, \cdot)$ is an open problem with limited theoretical results. Groeneboom and Wellner (1992) discussed the asymptotic of the behavior of the NPMLE in a version of the case 2 censoring model ($K_i = 2$). Moreover, Wellner (1995) studied the consistency when each subject gets exactly $k$ known examination times, and van der Vaart and Wellner (2000) proved the consistency of the maximum likelihood estimator of the mixed case interval censoring in the Hellinger distance; see also Schick and Yu (2000) and Song (2004).

When $\boldsymbol{\alpha}$ is unknown, the study of the asymptotic behavior of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, \cdot)$ and $\widehat{\boldsymbol{\alpha}}_n$ is even more challenging. The estimation of $\boldsymbol{\alpha}$ is coupled with the estimation of $\Lambda(\boldsymbol{\alpha}, \cdot)$. Therefore, unlike the Cox-type or general transformation model (Wellner and Zhang, 2007; Zeng et al., 2016), the limiting distribution of $\widehat{\boldsymbol{\alpha}}_n$ involves the limiting distribution and local behavior of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, \cdot)$ with respect to $(\boldsymbol{\alpha}, t)$ as well as the distribution of the frailty variable. On the other hand, the conditional estimating equation is constructed to avoid estimating the distribution of the frailty variable, which makes it different from the estimation of bundled parameters studied in Ding and Nan (2011). To the best of our knowledge, the limiting distributions of $\widehat{\boldsymbol{\alpha}}_n$ and $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, \cdot)$ remains an open problem.

Given the theoretical challenges, we consider making inferences about $\Lambda_0(t)$ and $\boldsymbol{\alpha}$ through a bootstrap procedure. The standard bootstrap variance estimator is reliable in problems with standard $n^{1/2}$-convergence rate, but is known to be inconsistent for NPMLE with

non-standard convergence rates in situations such as interval censored data (Abrevaya and Huang, 2005; Sen et al., 2010; Sen and Xu, 2015). Since the estimation of $\Lambda_0(t)$ was done by maximizing a working likelihood analogous to that in interval-censored data, the standard bootstrap estimate of $\widehat{\Lambda}_n(\boldsymbol{\alpha}, t)$ may suffer from inconsistency issues even when the true value $\boldsymbol{\alpha}$ is known; this would further lead to inconsistent estimation in the distribution of $\widehat{\boldsymbol{\alpha}}_n$. For this reason, we propose a model-based smoothed bootstrap procedure that provides a variance estimate with better agreement with the empirical one.

In particular, let $\widetilde{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, t)$ be a kernel-smoothed version of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, t)$. The smoothed bootstrap sampling procedure consists of two steps. First, a sample of the $n$ subjects is drawn with replacement from the original data. Second, for the $i$th subject in the sample, we keep the number of examinations $K_i^*$ and the examination times $t_{ij}^*$ $j = 1, \ldots, K_i^*$, but generate the panel counts $\{N_i^*(t_{ij}^*) - N_i^*(t_{i,j-1}^*); j = 1, \ldots, K_i^*\}$, from a working multinomial distribution with size $m_i^*$ and event probabilities proportional to $\widetilde{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, t_{ij}^*) - \widetilde{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, t_{ij-1}^*)$, $j = 1, \ldots, K_i^*$. The difference from the standard bootstrap is the second step. In the standard bootstrap sample, one subject may appear multiple times and all the appearances are the same as the observed data. In the smoothed bootstrap sample, the multiple appearances of the same subject may have different panel counts because they are independently regenerated from the fitted model. For each bootstrap sample, we apply our estimation procedure to obtain one draw of $\widehat{\boldsymbol{\alpha}}_n$ and $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, t)$. The empirical distributions of bootstrap replicates are then used to make inferences about $\boldsymbol{\alpha}$ and $\Lambda(t)$. In the simulation and data analysis, we considered the Nadaraya-Watson kernel regression with a Gaussian kernel and bandwidth determined by an unbiased cross-validation. More detailed specifications can be found in the Supplementary Materials.

The consistency of the standard bootstrap procedure depends on the limiting distribution of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, \cdot)$ and the consistency of $\widehat{\boldsymbol{\alpha}}_n$. For current status data, bootstrap consistency has

been explored in Sen and Xu (2015). For panel count data, this remains a challenging problem. In practice, the standard bootstrap procedure might be applicable for large sample sizes, but the model-based smoothed bootstrap procedure is generally recommended.

## 3. Simulation Study

Simulation studies were carried out to evaluate the performance of the proposed estimators. Two baseline functions were considered, $\lambda_0(t) = 2$ or $\lambda_0(t) = 2t$, for $t \in [0, \tau]$ with $\tau = 10$. For the $i$th subject, the covariates $X_{i1}$ and $X_{i2}$ were independently generated from the Bernoulli distribution with rate 0.5 and the uniform distribution over $[0, 1]$, respectively. The regression parameters were set at $\boldsymbol{\alpha} = (-1, -1)^\top$. The subject-specific frailty $Z_i$'s were generated from either a gamma distribution with mean 1 and variance 0.5 or a uniform distribution over $[0, 2]$, abbreviated by Gamma$(2, 2)$ and Uniform$(0, 2)$. Conditioning on $Z_i$, the recurrent event process was generated with inter-arrival times from either an exponential distribution or a uniform distribution first and then thinned so that Model 1. The exponential case results in a Poisson process on the individual level but the uniform case does not.

Depending on $Z_i$, the examination times were generated as follows. For $Z_i > 1$, $K_i$ was generated from a discrete uniform distribution on $\{1, \ldots, 8\}$ and the distinct examination times $t_{i1}, \ldots, t_{iK_i}$ were the order statistics of $K_i$ independent and identically distributed right truncated (by $\tau = 10$) exponential random variables with rate 2; for $Z_i \leqslant 1$, $K_i$ was generated from a discrete uniform distribution on $\{1, \ldots, 6\}$ and $t_{i1}, \ldots, t_{iK_i}$ were the order statistics of $K_i$ independent and identically distributed uniform random variable on $[0, 10]$. This design implies positive association between the underlying recurrent event process and the examination time process; subjects with $Z_i > 1$ have a higher event rate and tend to be examined more frequently than subjects with $Z_i \leqslant 1$. On average, the number of the recurrent events per subject ranged from 4 to 8 in all the configurations.

Three sample sizes were considered: $n = 50, 100, 200$. For variance estimation, the boot-

strap sample size was set to be 200 for both the standard bootstrap and the model-based smoothed bootstrap procedures. For each configuration, 1000 datasets were generated and analyzed. The computation task was demanding, and the SQUAREM implementation in the baseline hazard function estimation considerably reduced the running time; see Web Table 1 for a timing comparison for selected configurations.

Table 1 summarizes the results for the regression coefficient estimation based on 1000 replicates. The estimator appears to be unbiased in most scenarios. Noticeable bias (about 10%) only occurred in a couple of cases under $n = 50$ with event times generated from a non-Poisson process, which quickly diminishes as the sample size increases. For all scenarios, bootstrap standard error estimates from both procedures are reasonably close to the empirical standard errors, suggesting that the bootstrap estimator satisfactorily approximate the true variation for statistical inferences. For small sample ($n = 50$), the smoothed bootstrap standard errors appear to be a bit closer to the empirical standard errors and consequently, yield a coverage rate closer to the nominal level of 95% for the confidence intervals than the standard bootstrap standard errors. As expected, sample size $n = 200$ results in the best agreement between the bootstrap standard errors and the empirical standard errors, and between the empirical coverage rates and the nominal level of the confidence intervals.

Figure 1 presents the estimates and the pointwise 95% confidence intervals for the baseline cumulative rate function with $n = 50$. Since the baseline cumulative rate function is estimated under the transformed time scale, the baseline cumulative rate function can only be estimated between 0 and $\max_i(Y_i^*)$. This is reflected in Figure 1 where the average of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}, t)$ is almost indistinguishable from the truth for $t \in (0, 6)$, which covers the lower 98% of $Y_i^*$'s. Results with $n \in \{100, 200\}$ were similar and not reported.

In addition to sample size, the performance of the proposed estimator might depends on the strength and direction of the association between the underlying recurrent event

and examination time processes. In the Supplementary Material, we carried out additional simulations with different frailty distributions and examined scenarios where the recurrent event process and the examination time process are negatively correlated. In these settings, our estimator remains virtually unbiased, with bootstrap standard errors reasonably close to the empirical standard errors. Although the variance increases with the variance of the frailty distribution as expected, the empirical coverage rates are close to the nominal level in all scenarios. These results confirm the robustness of the proposed estimator.

## 4. Skin Cancer Chemoprevention Trial

In a double-blinded, placebo-controlled, randomized Phase III clinical trial (Bailey et al., 2010) conducted at the University of Wisconsin Comprehensive Cancer Center, the primary objective was to determine whether the application of difluoromethylornithine (DFMO) as a chemoprevention agent would lead to a significant reduction in the occurrence of two types of non-melanoma skin tumor: basal cell carcinomas (BCC) and squamous cell carcinomas (SCC). This study consisted of 290 patients with a history of skin cancer randomized into two groups: a treatment group with oral DFMO at a daily dose of 0.5 gram/m$^2$ and a placebo group. These patients were followed for 3 to 5 years depending on their entry time. Throughout the study, patients were scheduled to be examined every six months, but the scheduled times were followed only loosely instead of exactly. At each examination time, the number of newly developed skin tumors of each type were counted, measured and removed.

Of the 290 patients, 143 (49.3%) patients were in the DFMO group. The majority of the patients were male ($n$ = 174, 60%) and the age at enrollment ranged from 34 to 82 years with a median of 62 years. After the initial contact, the number of additional follow-up visits ranges from 0 to 16 with an average of 7.7. Figure 2(a) and Figure 2(b) show the tile plots for the two types of skin tumor counts observed at each visit. Each tile represents an examination time in days, with darker gray indicating larger count of new skin tumor

occurrences since the last visit. Although it is rare, a patient can develop both BCC or SCC tumors simultaneously. The figures indicate higher incidence in BCC tumor than in SCC tumor. The difference between the DFMO group and the placebo group appear to be small.

Table 2 summarizes the results of the data analysis based on three panel counts of skin tumors: the combined counts of two non-melanoma skin tumor (NMSC) types, the count of BCC, and the count of SCC. Four risk factors were considered as covariates: treatment group (1 = treatment, 0 = placebo), the number of prior non-melanoma skin tumor from diagnosis to randomization (ranges from 1 to 35, with mean 4.6), gender (1 = male, 0 = female), and age at enrollment (1 = age $\geqslant$ 65, 0 = otherwise). The estimated standard errors are obtained from the two bootstrap procedures, each with 500 bootstrap replicates. Gender, age and DFMO treatment did not seem to have any significant effect in reducing the recurrence either type of skin tumor or non-melanoma skin tumor. Controlling for other variables, the new skin tumor count was found to be significantly associated with the number of prior skin cancers. For every additional prior skin tumor, the time to a new BCC (or SCC) tumor development was estimated to shrink by a factor of $\exp(-0.155) \approx 0.856$ (or $\exp(-0.146) \approx 0.864$). These results are consistent with those in Li et al. (2011).

Figure 3 shows the average estimate and average pointwise 95% confidence intervals for $\Lambda(t)$ for the three outcome variables. Since the transformed time is inflated by positive coefficient estimates and large covariate values, we focus on the estimations in the time interval of $(0, 3600)$ days, where the right end was obtained by transforming the 98th percentile of the observed follow-up time by estimated coefficients with prior tumor count at its average and the other binary variables at zero. The two bootstrap procedures yielded very similar confidence intervals at earlier times, but at later times, the intervals from the smoothed bootstrap becomes noticeably narrower than those from the standard bootstrap for the

combined non-melanoma tumor and for the squamous cell carcinomas. This may be due to the inconsistency of the standard bootstrap.

To check the adequacy of the proposed model on the skin cancer data described in the main manuscript, we considered a graphical diagnosis based on Pearson type residuals of the observed counts $N_i(t_{ij})$'s conditioning on both the covariates $X_i$ and the frailty $Z_i$. It follows from Equation (3) that $E[m_i\Lambda_0^{-1}\{Y_i^*(\boldsymbol{\alpha})\} \mid Y_i, Z_i, \boldsymbol{X}_i] = Z_i$. Thus $E\{N_i(t) \mid Y_i, Z_i, \boldsymbol{X}_i\}$ can be approximated by $m_i\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}_n, t)/\widehat{\Lambda}_n\{\widehat{\boldsymbol{\alpha}}_n, Y_i^*(\widehat{\boldsymbol{\alpha}}_n)\}$. Under a working Poisson assumption, Pearson type residuals are obtained by standardizing the residuals $N_i(t) - E\{N_i(t) \mid Y_i, Z_i, \boldsymbol{X}_i\}$ with conditional variance $\mathrm{Var}\{N_i(t) \mid Y_i, Z_i, \boldsymbol{X}_i\} = E\{N_i(t) \mid Y_i, Z_i, \boldsymbol{X}_i\}$. Figure 4 presents the Pearson type residuals against the fitted value $E\{N_i(t) \mid Y_i, Z_i, \boldsymbol{X}_i\}$ for the three outcomes. The residuals are centered about zero and reveal no alarming patterns. The variance of the frailty was estimated as 7.8, 13.2, and 16.3 for the three outcomes, respectively, suggesting the necessity of accounting for the subject heterogeneity beyond the covariates; the heterogeneity level for SCC appears to be highes among the three.

## 5. Discussion

We considered a semiparametric accelerated mean model for panel count data under informative examination times. The AFT-type model offers an appealing alternative to the popular Cox-type models with covariate effects modifying the time scale of the cumulative mean function. In contrast to existing methods, our approach requires neither the strong Poisson-type assumption for the underlying recurrent event process nor a parametric assumption on the distribution of the unobserved frailty. The distribution of the examination time process is also left unspecified, thus allowing for an arbitrary association between the two processes. Consequently, the proposed method does not provide a direct characterization about the dependence between the two processes. The proposed method is most useful when the distributions of examination times and follow-up times are not of study interest.

When the covariate effects on the examination time process are of interest, a model similar to (1) may be imposed on the examination times. Specifically, let $O_i(t)$ be the number of examination times by time $t$ of subject $i$. Under conditional independence of $N_i(\cdot)$ and $O_i(\cdot)$ given $Z_i$ and $\boldsymbol{X}_i$, a joint scale-change model can be formulated by coupling Model (1) with $E\{\mathrm{d}O_i(t) \mid Z_i, \boldsymbol{X}_i\} = \nu(Z_i) r_0(te^{\boldsymbol{X}_i^\top \boldsymbol{\beta}}) e^{\boldsymbol{X}_i^\top \boldsymbol{\beta}} \mathrm{d}t$, where $\nu$ is an unspecified nonnegative function, $r_0(t)$ is an unspecified baseline rate function, and $\boldsymbol{\beta}$ is the regression coefficient vector. With such a joint model, our approach can still be applied directly to estimate $\boldsymbol{\alpha}$, while the method of Xu et al. (2017) can be used to estimate $\boldsymbol{\beta}$.

Diagnosis tools for the proposed model merit further investigation. The goodness-of-fit testing procedure for Cox-type rate function in Sun and Zhao (2013, Section 5.5.4) cannot be easily adapted to our setting because it requires the specification for the examination time process. Our graphical diagnosis based on Pearson residuals is only exploratory. A formal test procedure may be possible based on summaries of the Pearson residuals, with significance level assessed by bootstrap procedures. On a related issue, a general class of models that nests both the accelerated mean model and the Cox-type model would facilitate model selection. This class of model has been studied under noninformative censoring for univariate survival data (Chen and Jewell, 2001) and recurrent event data (Sun and Su, 2008). Extension to handle informative censoring to recurrent event data and panel count data is work in progress.



## 6. Supplementary Materials

Web Appendices and Tables referenced in Section 2.2, 2.3 and 3 are available with this paper at the Biometrics website on Wiley Online Library. An `R` package `spef` (Chiou et al., 2017) implementing the proposed method is available on the Comprehensive R Archive Network (R Core Team, 2017).

**Acknowledgement**

This research was partially supported by Harvard NeuroDiscovery Center and NIH T32NS048005 to Chiou, NSF SES1659328 and NSA H982301710308 to Xu, and NIH R01CA193888 to Huang.

**References**

Abrevaya, J. and Huang, J. (2005). On the bootstrap of the maximum score estimator. *Econometrica* **73,** 1175–1204.

Bailey, H. H., Kim, K., Verma, A. K., Sielaff, K., Larson, P. O., Snow, S., Lenaghan, T., Viner, J. L., Douglas, J., and Dreckschmidt, N. E. (2010). A randomized, double-blind, placebo-controlled phase 3 skin cancer prevention study of $\alpha$-difluoromethylornithine in subjects with previous history of skin cancer. *Cancer Prevention Research* **3,** 35–47.

Barzilai, J. and Borwein, J. M. (1988). Two-point step size gradient methods. *IMA Journal of Numerical Analysis* **8,** 141–148.

Buzkova, P. (2010). Panel count data regression with informative observation times. *The International Journal of Biostatistics* **6,** Article 30.

Buzkova, P. and Lumley, T. (2007). Longitudinal data analysis for generalized linear models with follow-up dependent on outcome-related variables. *Canadian Journal of Statistics* **35,** 485–500.

Chen, Y. Q. and Jewell, N. P. (2001). On a general class of semiparametric hazards regression models. *Biometrika* **88,** 687–702.

Chiou, S. H., Wang, X., and Yan, J. (2017). *spef: Semiparametric Estimating Functions*. R package version 1.0-5.

Ding, Y. and Nan, B. (2011). A sieve M-theorem for bundled parameters in semiparametric models, with application to the efficient estimation in a linear model for censored data.

*The Annals of Statistics* **39,** 3032–3061.

Ghosh, D. and Lin, D. Y. (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* **59,** 877–885.

Groeneboom, P. and Wellner, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*, volume 19 of *DMV Seminar*. Birkhäuser Verlag, Basel.

He, X., Tong, X., and Sun, J. (2009). Semiparametric analysis of panel count data with correlated observation and follow-up times. *Lifetime Data Analysis* **15,** 177–196.

Hu, X., Sun, J., and Wei, L.-J. (2003). Regression parameter estimation from panel counts. *Scandinavian Journal of Statistics* **30,** 25–43.

Hua, L. and Zhang, Y. (2012). Spline-based semiparametric projected generalized estimating equation method for panel count data. *Biostatistics* **13,** 440–454.

Huang, C.-Y., Wang, M.-C., and Zhang, Y. (2006). Analysing panel count data with informative observation times. *Biometrika* **93,** 763–775.

Kalbfleisch, J. and Lawless, J. F. (1985). The analysis of panel data under a markov assumption. *Journal of the American Statistical Association* **80,** 863–871.

Kim, Y.-J. (2006). Analysis of panel count data with dependent observation times. *Communications in Statistics-Simulation and Computation®* **35,** 983–990.

La Cruz, W., Martínez, J., and Raydan, M. (2006). Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. *Mathematics of Computation* **75,** 1429–1448.

Li, N., Park, D.-H., Sun, J., and Kim, K. (2011). Semiparametric transformation models for multivariate panel count data with dependent observation process. *Canadian Journal of Statistics* **39,** 458–474.

Li, N., Sun, L., and Sun, J. (2010). Semiparametric transformation models for panel count data with dependent observation process. *Statistics in Biosciences* **2,** 191–210.

Li, N., Zhao, H., and Sun, J. (2013). Semiparametric transformation models for panel count data with correlated observation and follow-up times. *Statistics in Medicine* **32,** 3039–3054.

Lin, D. Y., Wei, L. J., and Ying, Z. (1998). Accelerated failure time models for counting processes. *Biometrika* **85,** 605–618.

Lu, M., Zhang, Y., and Huang, J. (2009). Semiparametric estimation methods for panel count data using monotone B-splines. *Journal of the American Statistical Association* **104,** 1060–1070.

Ma, L. and Sundaram, R. (2016). Analysis of gap times based on panel count data with informative observation times and unknown start time. *Journal of the American Statistical Association* Forthcoming.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schick, A. and Yu, Q. (2000). Consistency of the GMLE with mixed case interval-censored data. *Scandinavian Journal of Statistics* **27,** 45–55.

Sen, B., Banerjee, M., and Woodroofe, M. (2010). Inconsistency of bootstrap: The Grenander estimator. *The Annals of Statistics* **38,** 1953–1977.

Sen, B. and Xu, G. (2015). Model based bootstrap methods for interval censored data. *Computational Statistics & Data Analysis* **81,** 121–129.

Song, S. (2004). Estimation with univariate "mixed case" interval censored data. *Statistics Sinica* **14,** 269–282.

Sun, J., Tong, X., and He, X. (2007). Regression analysis of panel count data with dependent observation times. *Biometrics* **63,** 1053–1059.

Sun, J. and Wei, L. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *Journal of the Royal Statistical Society: Series B* **62,**

293–302.

Sun, J. and Zhao, X. (2013). *Statistical Analysis of Panel Count Data*. Springer.

Sun, L. and Su, B. (2008). A class of accelerated means regression models for recurrent event data. *Lifetime Data Analysis* **14,** 357–375.

Thall, P. F. and Lachin, J. M. (1988). Analysis of recurrent events: Nonparametric methods for random-interval count data. *Journal of the American Statistical Association* **83,** 339–347.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society: Series B* **38,** 290–295.

van der Vaart, A. and Wellner, J. A. (2000). Preservation theorems for Glivenko-Cantelli and uniform Glivenko-Cantelli classes. In *High dimensional probability, II*, volume 47, pages 115–133. Birkhäuser Boston, Boston, MA.

Varadhan, R. (2014). *SQUAREM: Squared extrapolation methods for accelerating fixed-point iterations*. R package version 2014.8-1.

Varadhan, R. and Gilbert, P. (2009). BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function. *Journal of Statistical Software* **32,** 1–26.

Varadhan, R. and Roland, C. (2008). Simple and globally convergent methods for accelerating the convergence of any EM algorithm. *Scandinavian Journal of Statistics* **35,** 335–353.

Wang, M.-C., Qin, J., and Chiang, C.-T. (2001). Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association* **96,** 1057–1065.

Wang, X., Ma, S., and Yan, J. (2013). Augmented estimating equations for semiparametric panel count regression with informative observation times and censoring time. *Statistica*

*Sinica* **23,** 359–381.

Wellner, J. A. (1995). Interval censoring, case 2: alternative hypotheses. *Lecture Notes-Monograph Series* **27,** 271–291.

Wellner, J. A. and Zhang, Y. (2007). Two likelihood-based semiparametric estimation methods for panel count data with covariates. *The Annals of Statistics* **35,** 2106–2142.

Xu, G., Chiou, S. H., Huang, C.-Y., Wang, M.-C., and Yan, J. (2017). Joint scale-change models for recurrent events and failure time. *Journal of the American Statistical Association* **112,** 794–805.

Zeng, D. and Cai, J. (2010). A semiparametric additive rate model for recurrent events with an informative terminal event. *Biometrika* **97,** 699–712.

Zeng, D., Mao, L., and Lin, D. (2016). Maximum likelihood estimation for semiparametric transformation models with interval-censored data. *Biometrika* **103,** 253–271.

Zhang, Y. (2002). A semiparametric pseudolikelihood estimation method for panel count data. *Biometrika* **89,** 39–48.

Zhao, X. and Tong, X. (2011). Semiparametric regression analysis of panel count data with informative observation times. *Computational Statistics & Data Analysis* **55,** 291–300.

Zhao, X., Tong, X., and Sun, J. (2013). Robust estimation for panel count data with informative observation times. *Computational Statistics & Data Analysis* **57,** 33–40.

Zhou, J., Zhang, H., Sun, L., and Sun, J. (2017). Joint analysis of panel count data with an informative observation process and a dependent terminal event. *Lifetime Data Analysis* **23,** 560–584.

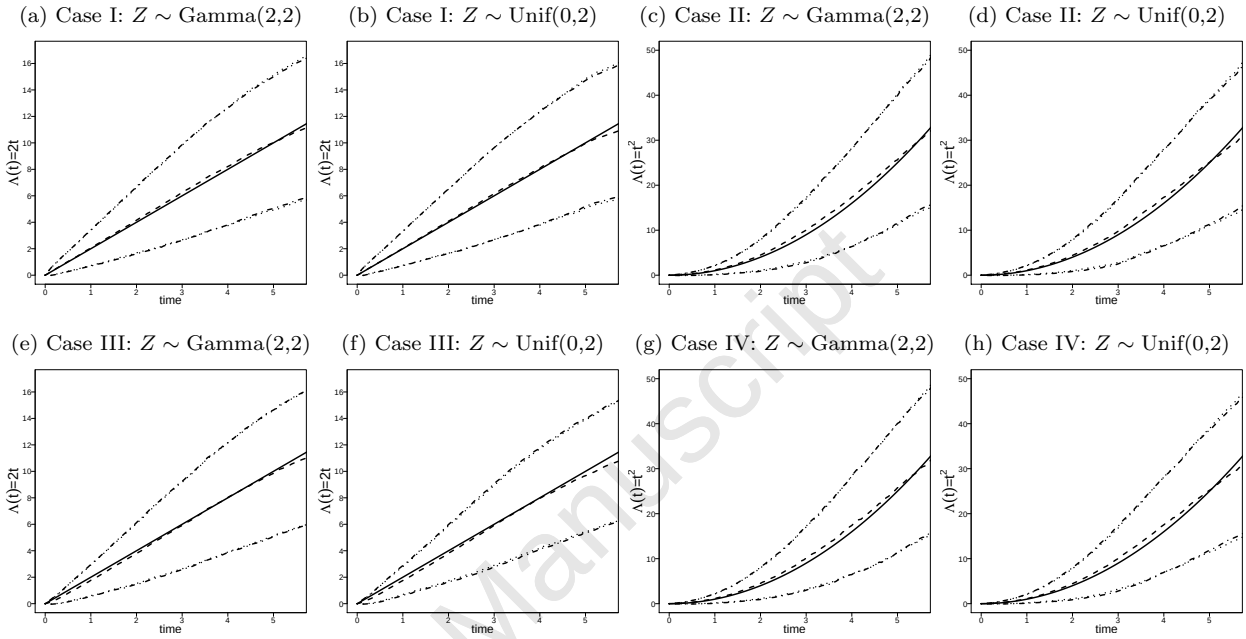


**Figure 1.** Plots of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}, t)$ with pointwise 95% confidence intervals for $n = 50$. Cases I–IV reflects the four combinations between the two choices of $\lambda_0(t)$ and whether the recurrent event process is a Poisson counting process; Case I: $\lambda_0(t) = 2$, Poisson process; Case II: $\lambda_0(t) = 2t$, Poisson process; Case III: $\lambda_0(t) = 2$, non-Poisson process; Case IV: $\lambda_0(t) = 2t$, non-Poisson process (—, true curve; - - -, empirical average; ·····, pointwise 95% standard bootstrap confidence intervals; -·-·-, pointwise 95% smoothed bootstrap confidence intervals).

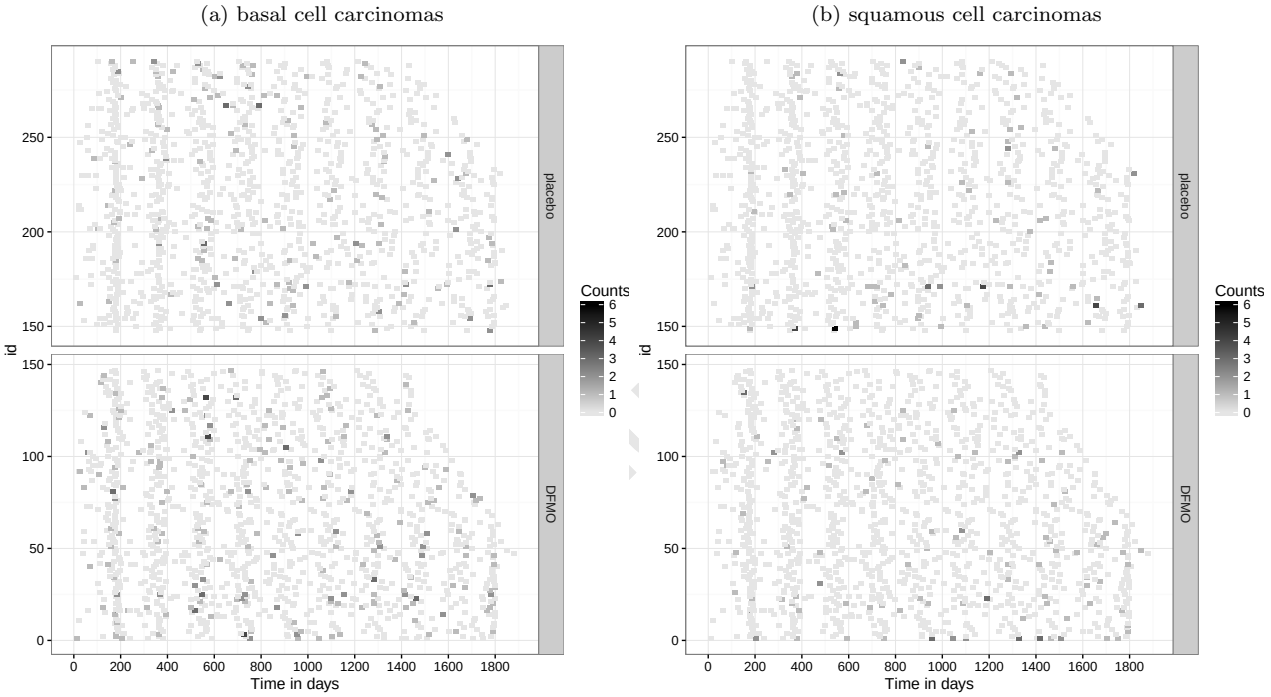


**Figure 2.** Tile plot of the skin cancer panel count. Each tile represents an examination time. Darker tiles represent larger numbers of tumor counts since the last visit.

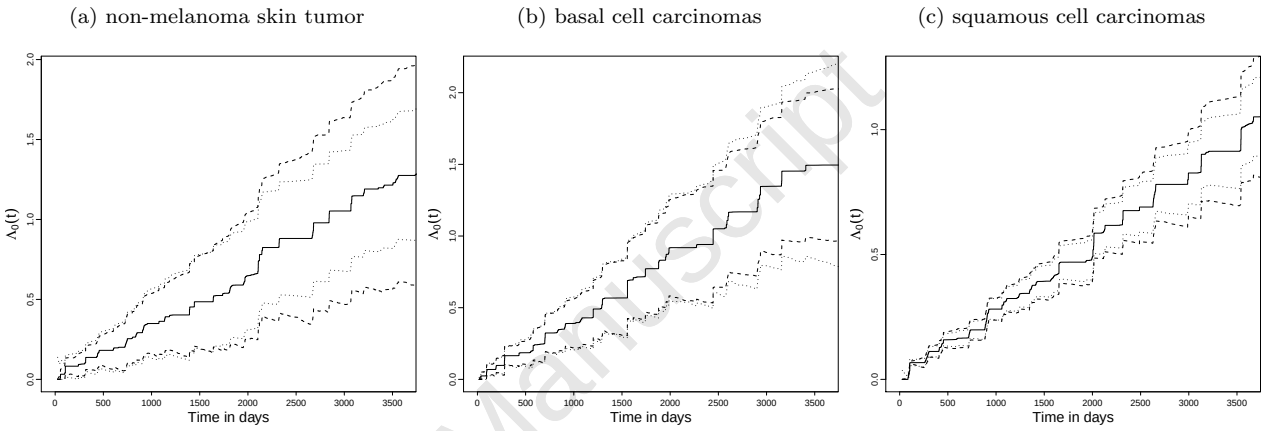


**Figure 3.** Plots of $\widehat{\Lambda}_n(\widehat{\boldsymbol{\alpha}}, t)$ for the Skin Cancer Chemoprevention Trial (- - -, pointwise 95% standard bootstrap confidence intervals; ·····, pointwise 95% smoothed bootstrap confidence intervals).

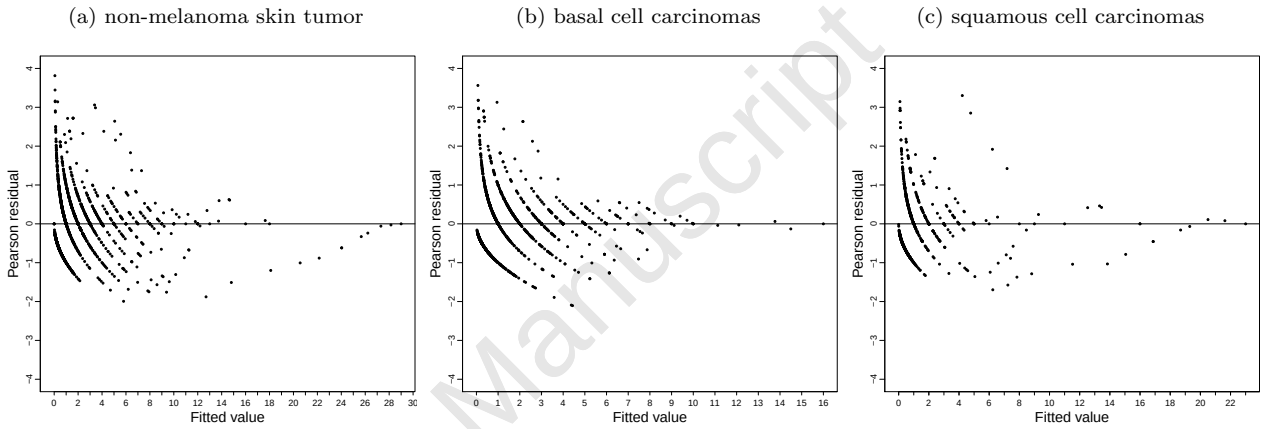


**Figure 4.** Pearson type residual plot for the Skin Cancer Chemoprevention Trial.

**Table 1**

*Summary of simulation data; ESE is the empirical standard error; ASE and ASE* are the average standard error based on the standard bootstrap and the smoothed bootstrap procedure, respectively; CP and CP* are the empirical coverage probability (%) based on the standard bootstrap and the smoothed bootstrap procedure, respectively. Cases I–IV reflects the four combinations between the two choices of $\lambda_0(t)$ and whether the recurrent event process is a Poisson counting process; Case I: $\lambda_0(t) = 2$, Poisson process; Case II: $\lambda_0(t) = 2t$, Poisson process; Case III: $\lambda_0(t) = 2$, non-Poisson process; Case IV: $\lambda_0(t) = 2t$, non-Poisson process.*

| | | $Z \sim \text{Gamma}(2,2)$ | | | | | | $Z \sim \text{Uniform}(0,2)$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| case | $\boldsymbol{\alpha}$ | bias | ESE | ASE | ASE* | CP | CP* | bias | ESE | ASE | ASE* | CP | CP* |
| | | | | | | $n = 50$ | | | | | | | |
| I | $\alpha_1$ | −0.009 | 0.315 | 0.309 | 0.314 | 93.6 | 94.2 | −0.027 | 0.298 | 0.302 | 0.308 | 95.7 | 95.7 |
| | $\alpha_2$ | −0.077 | 0.541 | 0.523 | 0.536 | 93.3 | 95.1 | −0.039 | 0.503 | 0.520 | 0.530 | 95.1 | 95.4 |
| II | $\alpha_1$ | −0.018 | 0.284 | 0.276 | 0.282 | 93.6 | 95.8 | −0.031 | 0.259 | 0.259 | 0.268 | 95.4 | 96.4 |
| | $\alpha_2$ | −0.083 | 0.492 | 0.470 | 0.482 | 92.2 | 94.1 | −0.050 | 0.461 | 0.446 | 0.457 | 92.6 | 93.4 |
| III | $\alpha_1$ | −0.082 | 0.226 | 0.215 | 0.224 | 91.6 | 93.5 | −0.091 | 0.237 | 0.212 | 0.225 | 92.9 | 95.3 |
| | $\alpha_2$ | −0.133 | 0.364 | 0.351 | 0.368 | 93.5 | 94.6 | −0.105 | 0.390 | 0.358 | 0.376 | 93.7 | 95.0 |
| IV | $\alpha_1$ | −0.088 | 0.206 | 0.210 | 0.218 | 93.0 | 95.1 | −0.085 | 0.214 | 0.208 | 0.218 | 93.9 | 95.5 |
| | $\alpha_2$ | −0.162 | 0.367 | 0.342 | 0.358 | 90.6 | 93.3 | −0.133 | 0.355 | 0.344 | 0.358 | 93.5 | 93.9 |
| | | | | | | $n = 100$ | | | | | | | |
| I | $\alpha_1$ | 0.002 | 0.213 | 0.216 | 0.217 | 94.1 | 94.7 | 0.010 | 0.199 | 0.208 | 0.210 | 96.2 | 96.4 |
| | $\alpha_2$ | −0.013 | 0.360 | 0.363 | 0.367 | 94.7 | 94.8 | −0.026 | 0.348 | 0.354 | 0.357 | 95.6 | 96.5 |
| II | $\alpha_1$ | 0.006 | 0.216 | 0.207 | 0.210 | 93.2 | 92.8 | 0.005 | 0.186 | 0.181 | 0.186 | 94.8 | 95.3 |
| | $\alpha_2$ | −0.019 | 0.358 | 0.343 | 0.348 | 93.6 | 93.5 | −0.028 | 0.312 | 0.312 | 0.316 | 95.1 | 95.8 |
| III | $\alpha_1$ | −0.048 | 0.154 | 0.151 | 0.154 | 93.1 | 93.9 | 0.018 | 0.152 | 0.161 | 0.158 | 96.0 | 94.1 |
| | $\alpha_2$ | −0.068 | 0.259 | 0.251 | 0.254 | 93.2 | 93.1 | −0.082 | 0.267 | 0.255 | 0.254 | 94.2 | 95.2 |
| IV | $\alpha_1$ | −0.057 | 0.151 | 0.147 | 0.148 | 91.8 | 92.6 | −0.064 | 0.156 | 0.147 | 0.149 | 91.7 | 94.8 |
| | $\alpha_2$ | −0.086 | 0.257 | 0.241 | 0.243 | 92.7 | 92.5 | −0.091 | 0.252 | 0.242 | 0.244 | 93.7 | 94.7 |
| | | | | | | $n = 200$ | | | | | | | |
| I | $\alpha_1$ | −0.003 | 0.160 | 0.157 | 0.157 | 96.4 | 94.4 | −0.003 | 0.145 | 0.150 | 0.143 | 95.9 | 95.7 |
| | $\alpha_2$ | −0.010 | 0.275 | 0.265 | 0.263 | 94.7 | 93.6 | −0.018 | 0.239 | 0.251 | 0.241 | 95.8 | 94.9 |
| II | $\alpha_1$ | −0.001 | 0.143 | 0.143 | 0.137 | 95.8 | 95.6 | 0.005 | 0.140 | 0.133 | 0.132 | 95.5 | 95.8 |
| | $\alpha_2$ | −0.013 | 0.249 | 0.244 | 0.234 | 95.8 | 94.8 | 0.006 | 0.235 | 0.228 | 0.222 | 95.3 | 95.2 |
| III | $\alpha_1$ | −0.029 | 0.122 | 0.116 | 0.117 | 94.6 | 94.2 | −0.010 | 0.113 | 0.117 | 0.109 | 93.3 | 94.2 |
| | $\alpha_2$ | −0.031 | 0.190 | 0.188 | 0.186 | 95.2 | 94.4 | −0.040 | 0.190 | 0.188 | 0.186 | 94.7 | 93.6 |
| IV | $\alpha_1$ | −0.041 | 0.111 | 0.111 | 0.105 | 93.7 | 93.1 | −0.035 | 0.114 | 0.112 | 0.111 | 93.4 | 93.8 |
| | $\alpha_2$ | −0.069 | 0.180 | 0.178 | 0.173 | 93.0 | 92.6 | −0.055 | 0.186 | 0.182 | 0.180 | 93.5 | 94.4 |

**Table 2**

*Summary of Skin Cancer Chemoprevention Trial data; BCC is basal cell carcinomas; SCC is squamous cell carcinomas; NMSC is the non-melanoma skin cancer including both BCC and SCC; DFMO is the difluoromethylornithine group; PE is the point estimate; SE is the standard error obtained from standard bootstrap; SE* is the standard error obtained from smoothed bootstrap.*

| | NMSC | | | BCC | | | SCC | | |
|---|---|---|---|---|---|---|---|---|---|
| Risk factor | PE | SE | SE* | PE | SE | SE* | PE | SE | SE* |
| DFMO | −0.003 | 0.149 | 0.216 | −0.063 | 0.176 | 0.186 | −0.003 | 0.124 | 0.198 |
| Prior tumor count | 0.152 | 0.024 | 0.048 | 0.155 | 0.027 | 0.026 | 0.146 | 0.030 | 0.032 |
| Male | 0.289 | 0.183 | 0.285 | 0.173 | 0.207 | 0.190 | 0.437 | 0.308 | 0.378 |
| 65 years or older | 0.088 | 0.134 | 0.196 | −0.226 | 0.196 | 0.179 | 0.595 | 0.339 | 0.405 |