



# On a simple estimation of the proportional odds model under right truncation

Peng Liu<sup>1</sup> · Kwun Chuen Gary Chan<sup>2</sup> · Ying Qing Chen<sup>3</sup>

Received: 15 January 2022 / Accepted: 1 December 2022 / Published online: 5 January 2023  
© The Author(s) 2023

## Abstract

Retrospective sampling can be useful in epidemiological research for its convenience to explore an etiological association. One particular retrospective sampling is that disease outcomes of the time-to-event type are collected subject to right truncation, along with other covariates of interest. For regression analysis of the right-truncated time-to-event data, the so-called proportional reverse-time hazards model has been proposed, but the interpretation of its regression parameters tends to be cumbersome, which has greatly hampered its application in practice. In this paper, we instead consider the proportional odds model, an appealing alternative to the popular proportional hazards model. Under the proportional odds model, there is an embedded relationship between the reverse-time hazard function and the usual hazard function. Building on this relationship, we provide a simple procedure to estimate the regression parameters in the proportional odds model for the right truncated data. Weighted estimations are also studied.

**Keywords** Biased sampling · Odds ratio · Reverse-time hazard function

**Mathematics Subject Classification** 62N02

## 1 Introduction

Truncation is common in survival analysis where the incomplete nature of the observations is due to a systematic biased selection process originated in the study design. Right truncated data arise naturally when an incubation period (i.e., the time between disease incidence and the onset of clinical symptoms) cannot be observed completely

---

✉ Peng Liu  
p.liu@kent.ac.uk

<sup>1</sup> School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7FS, UK

<sup>2</sup> Department of Biostatistics, University of Washington, Seattle, Washington 98195, USA

<sup>3</sup> Stanford Prevention Research Center, Palo Alto, California 94305, USA

in a retrospective study. In survival analysis, right truncation will lead to biased sampling in which shorter observations will be oversampled (Gürler 1996). For example, to study AIDS caused by blood transfusion (Lagakos et al. 1988), the incubation period is the time from a contaminated blood transfusion to the time when symptoms and signs of AIDS are first apparent. However, in those studies, the following-up period are usually limited. Therefore, only those developed AIDS before the end of study can be identified.

Many authors have studied right truncated data: Woodroffe (1985) and Wang et al. (1986) focused on the asymptotic properties the product limit estimator under random truncation. Keiding and Gill (1990) studied asymptotic properties of random left truncation estimator by a reparametrization of the left truncation model as a three-state Markov process. Lagakos et al. (1988) considered nonparametric estimation and inference of right truncated data by treating the process in reverse time, they showed that  $\lambda^B(t) = \lambda(\tau - t)$ , where  $\tau$  is the study duration,  $\lambda^B(t)$  and  $\lambda(t)$  are reverse-time hazard and forward-time hazard, respectively. The authors also discussed the implications and limitations of introducing reverse time hazard to analyze right truncated data. Gross (1992) further explained the necessity of reverse time hazard in the Cox model setting.

However, in most of the current literature, researchers study right truncated data in nonparametric setting, fairly few studied semiparametric models, among them, Kalbfleisch and Lawless (1989) formulating the Cox model on the reverse time hazard (or retro hazard, Lagakos et al. (1988); Keiding and Gill (1990)). For other related work on reverse time hazard, please refer to Gross (1992); Chen et al. (2004), among others.

In this paper, we study right truncated data under a semiparametric proportional odds model. Different from a proportional hazards model, the reverse-time hazard in proportional odds model has a simple log-linear relationship with the forward-time hazard, which leads to an intuitive estimator. While Sundaram (2009)'s method can also be adapted to proportional odds model for right truncated data, she focused on applying a reversed-time argument to an estimator for left truncated data. Our estimator, on the other hand, utilize a direct relationship between the reverse-time hazard, the forward-time hazard and the baseline odds function, so that we obtain a simpler estimator. Weighted functions are also being inserted into the estimating equation to obtain more efficient estimates.

The rest of the paper is organized as follows. Section 2 describes the inference procedure as well as asymptotic results, Sect. 3 shows simulation and real data results, Sect. 5 provides some discussion. Proof of theorems are left into the Appendix part.

## 2 Inference procedure

Assume that the failure time of interest  $T$  follows the semiparametric proportional odds model:

$$\log \left\{ \frac{1 - S(t | Z)}{S(t | Z)} \right\} = \alpha(t) + Z^\top \beta, \quad (1)$$

and the observed failure time is subject to a right truncation time variable  $R$ . The observed data is  $(T_i, R_i)$ ,  $i = 1, \dots, n$ , where  $T_i \leq R_i$ . Let  $\tau$  be the study duration, which is greater than  $\max\{T_1, T_2, \dots, T_n\}$ . An (observed) reverse-time sample,  $(T_i^*, R_i^*)$ ,  $i = 1, \dots, n$  can be constructed, where  $T^* = \tau - T$ ,  $R^* = \tau - R$ , so that  $T^*$  is left truncated by the variable  $R^*$ . Denote  $(\tilde{T}^*, \tilde{R}^*)$  as the reverse-time sample (potentially truncated). Then the hazard function of  $\tilde{T}^*$  is a quantity originated in  $\tau$  and counts backward in time. The reverse hazard and cumulative reverse hazard function of backward recurrence time is defined as

$$\lambda^B(t | Z) = \lim_{\Delta t \rightarrow 0} \frac{\Pr\{\tilde{T}^* \in (t - \Delta t, t] | \tilde{T}^* \leq t, Z\}}{\Delta t} = \frac{f(t | Z)}{F(t | Z)},$$

$$\Lambda^B(t | Z) = \int_t^\tau \lambda^B(s | Z) ds.$$

We would like to mention that a similar definition of the reverse hazard can also be found in Kalbfleisch and Lawless (1989) and Jiang (2011). Denote  $v(t) = \exp(\alpha(t))$ , and  $\lambda(t) = f(t)/S(t)$  as the forward-time hazard, then

$$\begin{aligned} \log \lambda(t | Z) - \log \lambda^B(t | Z) &= \alpha(t) + Z^\top \beta, \lambda^B(t | Z) \\ &= \frac{1}{\{1 + v(t) \exp(Z^\top \beta)\}v(t)} \frac{dv(t)}{dt}. \end{aligned}$$

Consider the counting process

$$N_i(t) = I(t \leq T_i \leq R_i), Y_i(t) = I(T_i \leq t \leq R_i),$$

and denote

$$M_i(t, \beta) = N_i(t) - \int_t^\tau Y_i(s) \frac{1}{\{\exp(Z_i^\top \beta)v(s) + 1\}v(s)} dv(s).$$

Then  $M_i(t, \beta)$  is a martingale with respect to the self-exciting (canonical) filtration (Keiding and Gill 1990; Stralkowska-Kominiak and Stute 2009) and

$$M_i(dt, \beta) = dN_i(t) + Y_i(t) \frac{1}{\{\exp(Z_i^\top \beta)v(t) + 1\}v(t)} dv(t). \tag{2}$$

Multiply both sides of (2) by  $\{\exp(Z_i^\top \beta)v(t) + 1\}$  and summing over  $n$  observations,

$$\begin{aligned} &\sum_{i=1}^n \{\exp(Z_i^\top \beta)v(t) + 1\}dN_i(t) + \sum_{i=1}^n Y_i(t) \frac{dv(t)}{v(t)} \\ &= \sum_{i=1}^n \{\exp(Z_i^\top \beta)v(t) + 1\}M_i(dt, \beta). \end{aligned} \tag{3}$$

Divide both left-hand side and right-hand side by  $\sum_{i=1}^n Y_i(t)$ , we obtain:

$$\frac{\sum_{i=1}^n \{\exp(Z_i^\top \beta)v(t) + 1\}dN_i(t)}{\sum_{i=1}^n Y_i(t)} + \frac{dv(t)}{v(t)} = \frac{\sum_{i=1}^n \{\exp(Z_i^\top \beta)v(t) + 1\}M_i(dt, \beta)}{\sum_{i=1}^n Y_i(t)}.$$

which is equivalent to:

$$\begin{aligned} v(t) \frac{\sum_{i=1}^n \exp(Z_i^\top \beta)dN_i(t)}{\sum_{i=1}^n Y_i(t)} + \frac{\sum_{i=1}^n dN_i(t)}{\sum_{i=1}^n Y_i(t)} + \frac{dv(t)}{v(t)} \\ = \sum_{i=1}^n \frac{\exp(Z_i^\top \beta)v(t) + 1}{\sum_{i=1}^n nY_i(t)} M_i(dt, \beta). \end{aligned} \tag{4}$$

Denote the left-hand side of (4) as:

$$U(\beta, dt) = \frac{dv(t)}{v(t)} + p_n(t)dt - q_n(t, \beta)v(t)dt,$$

where

$$p_n(t)dt = \frac{\sum_{i=1}^n dN_i(t)}{\sum_{j=1}^n Y_j(t)}, q_n(t, \beta)dt = -\frac{\sum_{i=1}^n \exp(Z_i^\top \beta)dN_i(t)}{\sum_{j=1}^n Y_j(t)}.$$

From standard counting process arguments (Anderson and Gill, 1982; Aalen10), we know that the stochastic integral with respect to the counting process martingale  $M_i(dt, \beta)$  is also a martingale, motivate by the following equation

$$E \left[ \frac{1}{n} U(\beta, dt) \right] = E \left[ \frac{1}{n} \sum_{i=1}^n \frac{\exp(Z_i^\top \beta)v(t) + 1}{\sum_{i=1}^n nY_i(t)} M_i(dt, \beta) \right].$$

We construct the following estimating equation

$$\frac{1}{n} U(\beta, dt) = 0. \tag{5}$$

Only  $v(t)$  is unknown in (5), let the estimate of  $v(t)$  be  $\hat{v}_n(t, \beta)$ . Denote

$$P_n(t) = \exp \left\{ \int_t^\tau \frac{\sum_{i=1}^n dN_i(s)}{\sum_{j=1}^n Y_j(s)} \right\}, Q_n(t, \beta) = \int_t^\tau \frac{\sum_{i=1}^n \exp(Z_i^\top \beta)dN_i(s)}{\sum_{j=1}^n Y_j(s)},$$

then

$$\hat{v}_n(t, \beta) = \frac{P_n(t)}{\int_t^\tau P_n(s)Q_n(ds, \beta)}. \tag{6}$$

Multiply (2) by  $Z_i \{ \exp(Z_i^\top \beta) v(t) + 1 \} / n$  and summing over  $n$  observations, we obtain

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n Z_i \left[ \{ \exp(Z_i^\top \beta) v(t) + 1 \} dN_i(t) + Y_i(t) \frac{dv(t)}{v(t)} \right] \\ & = \frac{1}{n} \sum_{i=1}^n Z_i \{ \exp(Z_i^\top \beta) v(t) + 1 \} M_i(dt, \beta). \end{aligned} \tag{7}$$

By virtue of the same idea of (5), take integration on both sides of (7), we can also construct another equation:

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau Z_i \left[ \{ \exp(Z_i^\top \beta) v(t) + 1 \} dN_i(t) + Y_i(t) \frac{dv(t)}{v(t)} \right] = 0 \tag{8}$$

Substituting (6) into (8), we can obtain the estimate of  $\beta$  by solving the following equation:

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau Z_i \left[ \{ \exp(Z_i^\top \beta) \hat{v}_n(t, \beta) + 1 \} dN_i(t) + Y_i(t) \frac{\hat{v}_n(dt, \beta)}{\hat{v}_n(t, \beta)} \right] = 0.$$

Moreover, since

$$\frac{\hat{v}_n(dt, \beta)}{\hat{v}_n(t, \beta)} = - \frac{\sum_{k=1}^n dN_k(t)}{\sum_{l=1}^n Y_l(t)} - \frac{\sum_{k=1}^n \exp(Z_k^\top \beta) dN_k(t)}{\sum_{l=1}^n Y_l(t)} \hat{v}_n(t, \beta),$$

then

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau \{ Z_i - \bar{Z}(t) \} \{ \exp(Z_i^\top \beta) \hat{v}_n(t, \beta) + 1 \} dN_i(t) = 0,$$

where

$$\bar{Z}(t) = \frac{\sum_{i=1}^n Z_i Y_i(t)}{\sum_{j=1}^n Y_j(t)}.$$

Finally, let

$$S_n(\beta) = \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{ Z_i - \bar{Z}(t) \} \{ \exp(Z_i^\top \beta) \hat{v}_n(t, \beta) + 1 \} dN_i(t), \tag{9}$$

and denote the solution of  $S_n(\beta) = 0$  be  $\hat{\beta}_n$ , we have the following theorem:

**Theorem 1** Under assumptions A1-A4 in the Appendix,  $\sqrt{n}(\hat{\beta}_n - \beta_0)$  converges weakly to a mean-zero normal distribution, with covariance matrix  $U^{-1}V(U^{-1})^\top$ , where  $V$  is the covariance matrix of  $\sqrt{n}S_n(\beta_0)$ ,  $U = \lim_{n \rightarrow \infty} \{\partial S_n(\beta) / \partial \beta\} |_{\beta = \beta_0}$ . The  $k$ th row of  $U$  is:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\} \times \left\{ Z_{ik} \exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + \exp(Z_i^\top \beta_0) \frac{\partial \hat{v}_n(t, \beta)}{\partial \beta_k} |_{\beta = \beta_0} \right\} dN_i(t).$$

**Remark** For proportional odds model with the normal logit link:

$$\log \left\{ \frac{S(t | Z)}{1 - S(t | Z)} \right\} = \alpha(t) + Z^\top \beta. \tag{10}$$

Define

$$\tilde{M}_i(t, \beta) = N_i(t) - \int_t^\tau Y_i(s) \frac{\exp(Z^\top \beta)}{1 + \exp(Z^\top \beta)v(s)} dv(s),$$

we claim that  $\tilde{M}_i(t, \beta)$  is a martingale. Recall that  $v(t) = \exp(\alpha(t))$ , following (10), we have

$$S(t|Z) = \frac{\exp(\alpha(t) + Z^\top \beta)}{1 + \exp(Z^\top \beta)v(t)},$$

as a result, we can obtain

$$f(t|Z) = \frac{\exp(Z^\top \beta)v'(t)}{(1 + \exp(Z^\top \beta)v(t))^2}, \quad F(t|Z) = \frac{1}{1 + \exp(Z^\top \beta)v(t)}.$$

Following the definition of reverse hazard in Sect. 2, we can write the reverse hazard as

$$\tilde{\lambda}^B(t|Z) = \frac{f(t|Z)}{F(t|Z)} = \frac{\exp(Z^\top \beta)v'(t)}{1 + \exp(Z^\top \beta)v(t)}.$$

From the general definition of martingale in Fleming and Harrington (1991) (pp. 25), we can easily show that  $\tilde{M}_i(t, \beta)$  is a martingale. While for model (1),

$$\lambda^B(t|Z) = \frac{v'(t)}{1 + \exp(Z^\top \beta)v(t)},$$

and  $N_i(t) - \int_t^\tau Y_i(s)\lambda^B(t|Z)dt$  is the martingale.

The corresponding estimating equation under model (10) has the following form

$$S_n^{(1)}(\beta) = \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t, \beta)\} \left\{ \exp(Z_i^\top \beta) \hat{v}_n(t, \beta) + 1 \right\} dN_i(t), \tag{11}$$

where

$$\bar{Z}(t, \beta) = \frac{\sum_{i=1}^n Z_i Y_i(t) \exp(Z_i^\top \beta)}{\sum_{j=1}^n Y_j(t) \exp(Z_j^\top \beta)}.$$

Equation (11) also can be used to estimate  $\beta$ , however, comparing with (9), (11) is more complicated and more computational intensive, while the derivative of (9) with respect to  $\beta$  can be easily obtained. As a result, (9) can be easily solved by the newton raphson algorithm. In the following simulations, we will use estimating equation (9).

In addition to the unweighted object function (9), weighted object function can also being included to obtain a class of weighted estimators of  $\beta_0$ . This procedure is often used to minimize the sandwich estimate as well as improve the efficiency. The weighted version of object function is

$$S_{n,w}(\beta) = \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t) \{Z_i - \bar{Z}(t)\} \left\{ \exp(Z_i^\top \beta) \hat{v}_n(t, \beta) + 1 \right\} dN_i(t) = 0, \tag{12}$$

here  $W_n(t)$  is a predictable weight function with respect to the canonical filtration which converges to a non-random function  $w(t)$ . One of the common used weight function is the Prentice-Wilcoxon type function  $W_{n1}(t) = \hat{S}_{LB}(t)$ , where  $\hat{S}_{LB}(\cdot)$  is the Lynden Bell estimate of the baseline survival function for right truncated failure time data. Denote the corresponding estimate of  $\beta$  as  $\hat{\beta}_{n,w}$ . Then we have the following theorem:

**Theorem 2** *Under the same assumptions as Theorem 1, when  $n \rightarrow \infty$ , for a prespecified weight function  $W_n(\cdot) \rightarrow w(\cdot)$ ,  $\sqrt{n}(\hat{\beta}_{n,w} - \beta_0)$  converges weakly to a mean-zero normal distribution, with covariance matrix  $U_w^{-1} V_w (U_w^{-1})^\top$ , where  $V_w$  is the covariance matrix of  $\sqrt{n}S_{n,w}(\beta_0)$ ,  $U_w = \lim_{n \rightarrow \infty} \{\partial S_{n,w}(\beta) / \partial \beta\} |_{\beta=\beta_0}$ . The  $k$ th row of  $U_w$  is:*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t) \{Z_i - \bar{Z}(t)\} \left[ Z_{ik} \exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + \exp(Z_i^\top \beta_0) \left\{ \partial \hat{v}_n(t, \beta) / \partial \beta_k \right\} |_{\beta=\beta_0} \right] dN_i(t).$$

Recently, many people considered problem of finding the optimal weight in a weighted estimating equation, including Chen and Cheng (2005); Chen and Wang (2000); Chen et al. (2012), among others. To achieve this goal, we only need to find the  $w(t)$  such that  $U_w(\beta_0)^{-1} V_w(\beta_0) U_w(\beta_0)^{-1}$  achieves the minimum. Since both the

empirical weight function  $W_n(t)$  and its limit  $w(t)$  do not rely on unknown parameter  $\beta_0$ , it is reasonable to set  $\beta_0 = 0$ . Another explanation for letting  $\beta_0 = 0$  is that it represents the baseline distribution. Therefore, let  $\beta_0 = 0$ , then we have:

$$\begin{aligned}
 U_w(\beta_0) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t) \{Z_i - \bar{Z}(t)\} Z_i \exp(Z_i^\top \beta_0) v(t) dN_i(t) \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t) \{Z_i - \bar{Z}(t)\}^{\otimes 2} \exp(Z_i^\top \beta_0) v(t) dN_i(t) \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t) \{Z_i - \bar{Z}(t)\}^{\otimes 2} Y_i(t) \frac{\exp(Z_i^\top \beta_0) v'(t)}{\exp(Z_i^\top \beta_0) v(t) + 1} dt \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t) \{Z_i - \bar{Z}(t)\}^{\otimes 2} Y_i(t) \frac{v'(t)}{v(t) + 1} dt. \tag{13}
 \end{aligned}$$

$$\begin{aligned}
 V_w(\beta_0) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t)^2 \{Z_i - \bar{Z}(t)\}^{\otimes 2} \\
 &\quad \times \left\{ \exp(Z_i^\top \beta_0) v(t) + 1 \right\}^2 Y_i(t) \frac{1}{\left\{ \exp(Z_i^\top \beta_0) v(t) + 1 \right\} v(t)} v'(t) dt \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t)^2 \{Z_i - \bar{Z}(t)\}^{\otimes 2} Y_i(t) \left\{ \exp(Z_i^\top \beta_0) v(t) + 1 \right\} \frac{v'(t)}{v(t)} dt \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau W_n(t)^2 \{Z_i - \bar{Z}(t)\}^{\otimes 2} Y_i(t) \{v(t) + 1\} \frac{v'(t)}{v(t)} dt \tag{14}
 \end{aligned}$$

Apply the Cauchy-Schwarz inequality to  $U_w(\beta_0)^{-1} V_w(\beta_0) U_w(\beta_0)^{-1}$  and let  $\beta_0 = 0$ , then it follows that the optimal weight is proportional to

$$w(t) = \frac{v(t)}{(v(t) + 1)^2} = S(t) \{1 - S(t)\}, \tag{15}$$

which minimize the variance of  $\hat{\beta}_n$ . Since when (15) holds, we have

$$\begin{aligned}
 U_w(\beta_0) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\}^{\otimes 2} Y_i(t) \frac{v(t)v'(t)}{(v(t) + 1)^3} dt, \\
 V_w(\beta_0) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\}^{\otimes 2} Y_i(t) \frac{v(t)v'(t)}{(v(t) + 1)^3} dt,
 \end{aligned}$$

which means when  $\beta_0 = 0$ , given  $w(t) = S(t)\{1 - S(t)\}$ , we have  $U_w(\beta_0)^{-1} V_w(\beta_0) U_w(\beta_0)^{-1}$  achieves the minimum value  $U_w(\beta_0)^{-1}$  (or equivalently  $V_w(\beta_0)^{-1}$ ).

In simulation, let  $W_{n2}(t) = \hat{S}_{LB}(t) \{1 - \hat{S}_{LB}(t)\}$ , the results are shown in Table 1, it can be seen that the weight  $W_{n2}(t)$  achieve the minimal variance among the three estimators.



**Table 1** Simulation results

n		$\beta_n^{(1)}$			$\beta_n^{(2)}$			Cov(%)
		Bias $\times 10^3$	SSE $\times 10^3$	SEE $\times 10^3$	Bias $\times 10^3$	SSE $\times 10^3$	SEE $\times 10^3$	
300	Unweight	31	329	342	8	319	338	97
	Prentice-Wilcoxon	30	249	277	18	268	278	95
	$W_{n,2}(t)$	23	227	258	8	254	261	93
	Shen et al. (2017)	-2	271	NA	39	274	NA	NA
400	Unweight	20	278	289	7	271	288	96
	Prentice-Wilcoxon	12	213	240	12	227	241	95
	$W_{n,2}(t)$	8	194	222	5	211	225	94
	Shen et al. (2017)	-9	210	NA	7	251	NA	NA
500	Unweight	14	247	254	5	247	255	96
	Prentice-Wilcoxon	8	188	214	7	205	215	95
	$W_{n,2}(t)$	-1	172	198	-3	188	200	94
	Shen et al. (2017)	-21	187	NA	9	235	NA	NA
600	Unweight	11	222	229	7	227	232	96
	Prentice-Wilcoxon	5	173	195	5	186	196	95
	$W_{n,2}(t)$	-4	155	180	-4	172	182	95
	Shen et al. (2017)	-20	164	NA	-5	180	NA	NA

SSE The sampling standard deviation, SEE The sampling standard error, Cov The empirical coverage of approximate 95% confidence intervals

### 3 Simulation and real data

We perform simulation studies to evaluate the finite sample properties of the proposed estimator. In simulation, let  $\alpha(t) = 3 \log t$ ,  $\beta_0 = (1, 0.5)^\top$ ,  $Z_1$  is a continuous variable follows a uniform distribution from 0 to 2,  $Z_2$  is a discrete variable follows a Bernoulli distribution with probability 0.5. The failure time variable is generated from model (1). The right truncation variable follows a uniform distribution from 0 to 4. This makes the truncation rate equals to 20%. For each simulation, 1000 datasets are generated, in each dataset, there are  $n$  observations,  $n = 300, 400, 500, 600$ , respectively.  $W_{n1}(t)$  and  $W_{n2}(t)$  are chosen as the weight functions in weighted estimating equations. As is shown in Table 1, three estimation equations yield unbiased estimates and the empirical coverage probability is around nominal level 95%, when weighted function is incorporated into the estimation equation, the efficiency is greatly improved, and the variance achieve minimal for  $W_{n2}(t)$  under three estimates.

As pointed out by one of the referees and the associate editor, Shen et al. (2017) also studied right truncated data under linear transformation models, and we know that when the error term in the linear transformation model follows logistic distribution (Fine et al., 1998), the model becomes the proportional odds model. Let

$$\begin{aligned} N_i^\dagger(t) &= I(\tau - T_i \leq t) = I(T_i \geq \tau - t), \\ Y_i^\dagger(t) &= I(\tau - R_i \leq t \leq \tau - T_i) = I(T_i \leq \tau - t \leq R_i), \end{aligned}$$

then the estimating equations (3) and (4) in Shen et al. (2017) can be written as

$$\begin{aligned} &U(\beta, \alpha(\tau - t)) \\ &= \sum_{i=1}^n \int_{-\infty}^{\tau} Z_i \left[ dN_i(t) - Y_i(t) d \left( \log \frac{\exp(Z_i^\top \beta + \alpha(\tau - t))}{1 + \exp(Z_i^\top \beta + \alpha(\tau - t))} \right) \right] = 0, \\ &\sum_{i=1}^n \left[ dN_i(t) - Y_i(t) d \left( \log \frac{\exp(Z_i^\top \beta + \alpha(\tau - t))}{1 + \exp(Z_i^\top \beta + \alpha(\tau - t))} \right) \right] = 0. \end{aligned}$$

We recognize that Shen et al. (2017)'s methodology is general and works for all the linear transformation models, including the proportional odds model. However, our approach will be more convenient compared with Shen et al. (2017)'s under the proportional odds model, since our approach has a simpler form, and the estimation of the intercept  $\alpha(t)$  can be done beforehand and plugged in the final estimating equation, while Shen et al. (2017) can not achieve this and their estimation produce involves a complicated iteration which increases the risk of non-convergence. Besides, Shen et al. (2017) only deal with the reverse time but not the reverse hazard function, and we utilize the relationship between the reverse hazard function and the forward-time hazard function and produced a more intuitive estimator.

We conduct simulations for Shen et al. (2017)'s method and the results are reported in Table 1. The code was obtained from the authors via personal communication. However, one of the authors, Prof. Pao-Sheng Shen mentioned that they were unable to calculate the asymptotic variance and coverage probabilities, the existing results

in their paper contain some errors, and their current code only consists of bias and standard error. As a result, we only report bias and standard error of Shen et al. (2017)'s method. All the simulations were conducted under the same model as ours. We also want to mention that we found the computation speed is very slow for Shen et al. (2017)'s method, though asymptotic variance and coverage probability were not calculated, their method is still more than 3 times slower than ours under the same model setting and the sample size. The SSE of Shen et al. (2017)'s method is smaller than our unweighted estimator, but is bigger than the two weighted estimators. For the second approach in their paper, i.e. the conditional maximum-likelihood approach, since the bias is large, we did not perform further comparisons here. We would like to mention that the large bias of the conditional maximum-likelihood approach is also confirmed in Vakulenko-Lagun et al. (2020).

As suggested by one of the reviewers, we also perform simulations without accounting for the truncation, and the results are shown in Table 2. We choose the truncation distribution as uniform distributions from 0 to 4, 2 and 1, respectively, which corresponds to 20% truncation rate (mild truncation), 40% truncation rate (moderate truncation) as well as 70% truncation rate (heavy truncation). As we can see from Table 2, all the estimators are biased, and a larger truncation rate will lead to a bigger bias and variance, though for the same truncation, variances will decrease when the sample sizes increase. These results also coincide with Table 2.1 (pp. 20) in Rennert (2018) and Table 1 in Rennert and Xie (2018), though the two articles deal with the doubly truncated data under the Cox model.

To better illustrate how to employ the proposed method in real situation, we analyze the Centers for Disease Control's blood-transfusion data, this data was used by Kalbfleisch and Lawless (1989) and Wang (1989). The data include 494 cases reported to the Center of Disease Control prior January, 1, 1987, and diagnosed before July, 1, 1986. Only 295 of the 494 has consistent data, and they got infection by a single blood transfusion or a short series of transfusions, analyse is restricted to this subset. We obtain the raw observation data via personal communication, Thomas Peterman, Centers for Disease Control and Prevention. The data contains three variables:  $T$  is the time from blood transfusion to the diagnosis of AIDS (in months),  $R$  is the time from blood transfusion to the end of the study (July, 1986, in months), Age is the age of the person when transfusing blood (in years). Comparing the data with Kalbfleisch and Lawless (1989)'s as well as Wang (1989)'s, the observation ( $X=16$ ,  $T=33$ , Age=34) cannot be found in the raw data, thus is being deleted and the final sample size is 294, and a few fractions of the data are also corrected because these entries are not correct compared to the raw data.

We apply the proposed method to this data and treat Age as the covariate in regression. In Wang (1989)'s paper, the data are categorized into three age groups: 'children' aged 1-4, 'adults' aged 5-59, and 'elderly patients' aged 60 and older because of different patterns of survivorship, the survivor behaviour of groups 'adults' and 'elderly patients' are similar except for the right tail while there is an evident distinction compared with 'children', in current analysis, we delete the data from 'children', and focus on a combined sample of 'adults' and 'elderly patients' with a sample size equal to 260. Finally, the range of  $T$  is from 0 to 89, and the range of  $R$  is from 0 to 99. For all  $i \in \{1, \dots, 260\}$ , we have  $T_i \leq R_i$ . As a result, our dataset will not have the

**Table 2** Simulation results when ignoring truncation

n	Truncation	$\hat{\beta}_n^{(1)}$		$\hat{\beta}_n^{(2)}$		Cov(%)	SEE×10 <sup>3</sup>	SSE×10 <sup>3</sup>	Cov(%)	SEE×10 <sup>3</sup>	SSE×10 <sup>3</sup>
		Bias×10 <sup>3</sup>	SSE×10 <sup>3</sup>	Bias×10 <sup>3</sup>	SSE×10 <sup>3</sup>						
300	Mild	-39	284	-35	267	95	309	267	95	310	267
300	Moderate	-134	307	-74	297	86	305	297	86	308	297
300	Heavy	-306	555	-106	555	91	629	555	91	601	555
400	Mild	-67	229	-46	225	95	260	225	95	263	225
400	Moderate	-128	254	-71	265	89	261	265	89	264	265
400	Heavy	-379	245	-157	236	70	251	236	70	261	236
500	Mild	-66	200	-26	222	96	229	222	96	235	222
500	Moderate	-125	215	-72	227	91	230	227	91	234	227
500	Heavy	-398	215	-170	224	55	222	224	55	231	224
600	Mild	-53	173	-20	206	96	209	206	96	214	206
600	Moderate	-129	184	-72	213	93	208	213	93	212	213
600	Heavy	-406	223	-165	202	45	202	202	45	210	202

SSE, the sampling standard deviation; SEE, the sampling standard error; Cov, the empirical coverage of approximate 95% confidence intervals

**Table 3** Age effect for blood transfusion data

	Age	SSE
unweighted	-0.0128	0.0153
Prentice-Wilcoxon	-0.0120	0.0143
$W_{n2}(t)$	-0.0122	0.0122
Shen et al. (2017)	-0.0125	0.0150

identifiability issue as mentioned in Seaman et al. (2022). We also applied Shen et al. (2017)'s method and the result is similar. All the results are shown in Table 3, where the weights are chosen as  $W_{n1}(t)$  and  $W_{n2}(t)$ , the estimated parameter between unweighted and weighted estimation equation does not show much difference, but the variance is reduced when weights are considered. In both situations mentioned above, Age has a very weakening positive effect on the odds ratio, but the effect is not significant.

## 4 Discussion

Directly consider the right truncated data in normal time order can be failed because 'at risk' process is not adapt to the history of the process (Gross 1992). Retro hazard solves this problem which transform right truncated data to left truncated in reverse time (Woodroffe 1985). Statistical modelling is even more flexible by incorporating the nature structure of proportional odds model. The usual form of proportional odds model can also be utilized but the theoretical and computational burden for the estimator will be increased, employ (1) can substantially improve the situation.

**Acknowledgements** We thank Thomas Peterman, Centers for Disease Control and Prevention provided us the CDC blood transfusion data. We also thank for Pao-sheng Shen, Tunghai University provided us the code for Shen et al. (2017) and Vakulenko-Lagun Bella, University of Haifa discussed the simulation results of Vakulenko-Lagun et al. (2020).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix 1

Assumptions:

- A1:  $\beta_0 \in \mathbb{R}^p$  is the interior point of a compact set  $\mathcal{B}$ .
- A2:  $Z$  is a bounded process.
- A3:  $V(\beta_0)$  is non-negative.

A4:  $f(t)$  is continuous.

Assumption A1 is also used by Chen et al. (2012), A2 is a standard assumption to ensure martingale properties holds (Fleming and Harrington 1991), A3 is also a standard assumption to avoid theoretical discussion, it is also being used in Huang and Qin (2013), A4 is being used in prove the martingale representation of  $\hat{v}_n(t, \beta_0) - v(t)$ . Besides that, we also need an condition to ensure that the truncated distribution to be correctly identified, let  $F(\cdot)$  and  $G(t)$  be the distribution function of  $T$  and  $R$ , define  $(a_F, b_F)$  and  $(a_G, b_G)$  be the support of  $F(\cdot)$  and  $G(\cdot)$  of  $T$  and  $R$  under the meaning that  $a_W = \inf\{x : W(x) > 0\}, b_W = \sup\{x : W(x) < 1\}$ , where  $W$  is a distribution function. Under right truncation, actually, only conditional distribution  $P(T \leq x|T \leq b_G)$  and  $P(R \leq x|R \geq a_F)$  can be estimated, thus we assume  $a_F = a_G = 0, b_R = \infty$ , so that the conditional distribution will be the actual distribution of  $T$  and  $R$ , we also assume  $P(T \leq Y) = \alpha > 0$  to ensure that there exist observations satisfy our condition, similar assumption and discussion also appeared in Woodroffe (1985); Wang (1989), and Sundaram (2009), among others.

**Proof of Theorem 1** To prove the Theorem 1, the first step is to derive the martingale representation of  $\hat{S}_n(\beta_0)$ . To do this, we need the martingale representation of  $\hat{v}_n(t, \beta_0) - v_0(t)$ . Notice that

$$\begin{aligned} & \sum_{i=1}^n \{\exp(Z_i^\top \beta_0)v_0(t) + 1\}dN_i(t) + \sum_{i=1}^n Y_i(t) \frac{dv_0(t)}{v_0(t)} \\ &= \sum_{i=1}^n \{\exp(Z_i^\top \beta_0)v_0(t) + 1\}M_i(dt, \beta_0) \end{aligned} \tag{16}$$

$$\sum_{i=1}^n \{\exp(Z_i^\top \beta_0)\hat{v}_n(t, \beta_0) + 1\}dN_i(t) + \sum_{i=1}^n Y_i(t) \frac{\hat{v}_n(dt, \beta_0)}{dt} \frac{1}{\hat{v}_n(t, \beta_0)} = 0. \tag{17}$$

Denote  $w_0(t) = 1/v_0(t)$  and  $\hat{w}_n(t, \beta) = 1/\hat{v}_n(t, \beta)$ , then (17) and (16) becomes:

$$\begin{aligned} & \sum_{i=1}^n \{\exp(Z_i^\top \beta_0) + w_0(t)\}dN_i(t) - \sum_{i=1}^n Y_i(t)dw_0(t) \\ &= \sum_{i=1}^n \{\exp(Z_i^\top \beta_0) + w_0(t)\}M_i(dt, \beta_0), \end{aligned} \tag{18}$$

$$\sum_{i=1}^n \{\exp(Z_i^\top \beta_0) + \hat{w}_n(t, \beta_0)\}dN_i(t) - \sum_{i=1}^n Y_i(t)\hat{w}_n(dt, \beta_0) = 0. \tag{19}$$

(19)-(18) and divide both side by  $-\sum_{i=1}^n Y_i(t)$ :

$$\frac{\partial\{\hat{w}_n(t, \beta_0) - w_0(t)\}}{\partial t} - p_n(t)dt\{\hat{w}_n(t, \beta_0) - w_0(t)\}$$

$$= \frac{\sum_{i=1}^n \{\exp(Z_i^\top \beta_0) + w_0(t)\} M_i(dt, \beta_0)}{\sum_{i=1}^n Y_i(t)}.$$

Then

$$\hat{w}_n(t, \beta_0) - w_0(t) = \frac{1}{P_n(t)} \sum_{i=1}^n \int_t^\tau P_n(s) \frac{\{\exp(Z_i^\top \beta_0) + w_0(s)\}}{\sum_{j=1}^n Y_j(s)} M_i(ds, \beta_0).$$

In the interval  $(0, \tau)$ , since  $0 < v_0(t) < \infty$ , by delta method,

$$\begin{aligned} \hat{v}_n(t, \beta_0) - v_0(t) &= -\frac{1}{w_0^2(t)} \{\hat{w}_n(t, \beta_0) - w_0(t)\} \\ &= -\frac{v_0^2(t)}{P_n(t)} \sum_{i=1}^n \int_t^\tau P_n(s) \frac{v_0(s) \exp(Z_i^\top \beta_0) + 1}{\sum_{j=1}^n Y_j(s) v_0(s)} M_i(ds, \beta_0). \end{aligned} \tag{20}$$

At the point 0, (20) holds without condition because  $\hat{v}_n(0, \beta) = v_0(t) = 0$ . At the point  $\tau$ , if denote  $0 \times \infty = 0$ , then (20) also holds.

By using (20), for  $S_n(\beta_0)$ :

$$\begin{aligned} S_n(\beta_0) &= \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\} \left\{ \exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + 1 \right\} dN_i(t) \\ &= \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\} \left\{ \exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + 1 \right\} M_i(dt, \beta_0) \\ &\quad - \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\} Y_i(t) \frac{\exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + 1}{\exp(Z_i^\top \beta_0) v_0^2(t) + v_0(t)} dv_0(t) \\ &= \text{I} + \text{II}. \end{aligned}$$

In the following, we will show that the second part can also be represented as a summation of integral with respect to martingale.

$$\begin{aligned} \text{II} &= -\frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\} Y_i(t) \left\{ \frac{\exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + 1 - \exp(Z_i^\top \beta_0) v_0(t) - 1}{\exp(Z_i^\top \beta_0) v_0^2(t) + v_0(t)} \right. \\ &\quad \left. + \frac{1}{v_0(t)} \right\} dv_0(t) \\ &= -\frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t)\} Y_i(t) \frac{\exp(Z_i^\top \beta_0)}{\exp(Z_i^\top \beta_0) v_0^2(t) + v_0(t)} \{\hat{v}_n(t, \beta_0) - v_0(t)\} dv_0(t). \end{aligned} \tag{21}$$

Substitute (20) into (21) and change the integration order, then

$$\begin{aligned} \Pi &= \frac{1}{n} \sum_{j=1}^n \int_0^\tau \left\{ P_n(t) \frac{\exp(Z_j^\top \beta_0) v_0(t) + 1}{\sum_{k=1}^n Y_k(t) v_0(t)} \right. \\ &\quad \left. \sum_{i=1}^n \int_0^t \{Z_i - \bar{Z}(s)\} Y_i(s) \frac{\exp(Z_i^\top \beta_0) v_0(s)}{\exp(Z_i^\top \beta_0) v_0(s) + 1} \frac{1}{P_n(s)} d v_0(s) \right\} M_j(dt, \beta_0). \end{aligned}$$

Denote

$$\begin{aligned} \xi_i(t, \beta_0) &= \{Z_i - \bar{Z}(t)\} \left\{ \exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + 1 \right\} \\ &\quad + P_n(t) \frac{\exp(Z_i^\top \beta_0) v_0(t) + 1}{\sum_{k=1}^n Y_k(t) v_0(t)} \\ &\quad \times \sum_{j=1}^n \int_0^t \{Z_j - \bar{Z}(s)\} Y_j(s) \frac{\exp(Z_j^\top \beta_0) v_0(s)}{\exp(Z_j^\top \beta_0) v_0(s) + 1} \frac{1}{P_n(s)} d v_0(s). \end{aligned}$$

Then the martingale representation of  $S_n(\beta_0)$  is

$$S_n(\beta_0) = \frac{1}{n} \sum_{i=1}^n \int_0^\tau \xi_i(t, \beta_0) M_i(dt, \beta_0). \tag{22}$$

Through (22), it is obvious to prove that  $S_n(\beta_0)$  converges to zero 0 in probability by the weak law of large numbers.

Let

$$\mu(t) = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n Y_i(t) Z_i}{\sum_{j=1}^n Y_j(t)}, \quad v(t, \beta) = \lim_{n \rightarrow \infty} \hat{v}_n(t, \beta).$$

Denote

$$s_n(\beta) = \frac{1}{n} \sum_{i=1}^n \int_0^\tau \{Z_i - \mu(t)\} \left\{ \exp(Z_i^\top \beta) \hat{v}_n(t, \beta) - \exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) \right\} dN_i(t).$$

The derivative of  $S_n(\beta)$  and  $s_n(\beta)$  are

$$\begin{aligned} S'_n(\beta) &= \frac{1}{n} \sum_{i=1}^n \{Z_i - \bar{Z}(t)\} \left\{ \exp(Z_i^\top \beta) Z_i \hat{v}_n(t, \beta) + \exp(Z_i^\top \beta) \{\partial \hat{v}_n(t, \beta) / \partial \beta\} \right\} dN_i(t), \\ s'_n(\beta) &= \frac{1}{n} \sum_{i=1}^n \{Z_i - \mu(t)\} \left\{ \exp(Z_i^\top \beta) Z_i \hat{v}_n(t, \beta) + \exp(Z_i^\top \beta) \{\partial \hat{v}_n(t, \beta) / \partial \beta\} \right\} dN_i(t). \end{aligned}$$

Notice  $s_n(\beta_0) = 0$ . Assume that there exists  $\varepsilon > 0$  such that (A5):  $P\{|Z_i - \mu(t)| > \varepsilon, i = 1, 2, \dots, n\} > 0$ , which means covariate can not be identical for all individuals.



Together with the assumption (A6):

$$| E \left\{ \exp(Z^\top \beta_0) Z \hat{v}_n(t, \beta_0) \right\} + E \left[ \exp(Z^\top \beta_0) \{ \partial \hat{v}_n(t, \beta) / \partial \beta \} |_{\beta=\beta_0} \right] | > 0.$$

we have  $| \lim_n s'_n(\beta_0) | > 0$ . Without loss of generality, let  $\lim_n s'_n(\beta_0) > 0$ , then there exist a neighborhood of  $\beta_0$  such that  $s_n(\beta)$  is strictly increasing. Further notice that  $S_n(\beta) = s_n(\beta) + o_p(1)$ ,  $S'_n(\beta) = s'_n(\beta) + o_p(1)$ , then  $S_n(\beta)$  is strictly increasing in a neighborhood of  $\beta_0$ , thus prove the consistency of  $\hat{\beta}_n$ .

By martingale central limit theorem, the variance of  $S_n(\beta_0)$  is

$$\begin{aligned} V(\beta_0) &= \lim_{n \rightarrow \infty} V_n = \lim_{n \rightarrow \infty} \langle n^{-1/2} S_n(\beta_0), n^{-1/2} S_n(\beta_0) \rangle (\tau) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \int_0^\tau \sum_{i=1}^n \xi_i(t, \beta_0)^{\otimes 2} d \int_t^\tau -Y_i(s) \{ \exp(Z_i^\top \beta_0) v^2(s) + v(s) \}^{-1} dv(s) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \int_0^\tau \sum_{i=1}^n \xi_i(t, \beta_0)^{\otimes 2} Y_i(t) \{ \exp(Z_i^\top \beta_0) v^2(t) + v(t) \}^{-1} dv(t). \end{aligned}$$

Further using the delta method will complete the proof of Theorem 1. □

**Proof of Theorem 2** Since Theorem 1 and 2 are quite similar, in this part, we will omit the proof detail and only give the detailed expression of  $V_w$ .

$$V_w(\beta_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \int_0^\tau \sum_{i=1}^n \xi_{i,w}(t, \beta_0)^{\otimes 2} Y_i(t) \{ \exp(Z_i^\top \beta_0) v^2(t) + v(t) \}^{-1} dv(t).$$

where

$$\begin{aligned} \xi_{i,w}(t, \beta_0) &= W_n(t) \{ Z_i - \bar{Z}(t) \} \left\{ \exp(Z_i^\top \beta_0) \hat{v}_n(t, \beta_0) + 1 \right\} \\ &\quad + P_n(t) \frac{\exp(Z_i^\top \beta_0) v_0(t) + 1}{\sum_{k=1}^n Y_k(t) v_0(t)} \\ &\quad \times \sum_{j=1}^n \int_0^t W_n(s) \{ Z_j - \bar{Z}(s) \} Y_j(s) \frac{\exp(Z_j^\top \beta_0) v_0(s)}{\exp(Z_j^\top \beta_0) v_0(s) + 1} P_n(s)^{-1} dv_0(s). \end{aligned}$$

□

## References

Aalen OO, Andersen PK, Borgan Ø, Gill RD, Keiding N (2010) History of applications of martingales in survival analysis. arxiv preprint [arXiv:1003.0188](https://arxiv.org/abs/1003.0188)  
 Andersen PK, Gill RD (1982) Cox’s regression model for counting processes: a large sample study. Ann Statist 10:1100–20

- Chen YQ, Wang M-C (2000) Analysis of accelerated hazards models. *J Am Statist Assoc* 95:608–18
- Chen YQ, Wang M-C, Huang Y (2004) Semiparametric regression analysis on longitudinal pattern of recurrent gap times. *Biostatistics* 5:277–90
- Chen YQ, Cheng S (2005) Semiparametric regression analysis of mean residual life with censored survival data. *Biometrika* 92:19–29
- Chen YQ, Hu N, Musoke P, Zhao LP (2012) Estimating regression parameters in an extended proportional odds model. *J Am Statist Assoc* 107:318–30
- Fine JP, Ying Z, Wei LG (2012) On the linear transformation model for censored data. *Biometrika* 85:980–6
- Fleming TR, Harrington DP (1991) Counting processes and survival analysis. John Wiley, New York
- Gross ST (1992) Regression models for truncated survival data. *Scand J Stat* 19:193–213
- Gürler Ü (1996) Bivariate estimation with right-truncated data. *J Am Statist Assoc* 91:1152–65
- Huang C-Y, Qin J (2013) Semiparametric estimation for the additive hazards model with left-truncated and right-censored data. *Biometrika* 100:877–88
- Jiang Y (2011) Estimation of hazard function for right truncated data. Georgia State University, Thesis
- Kalbfleisch JD, Lawless JF (1989) Inferences based on retrospective ascertainment: an analysis of the data on transfusion-related AIDS. *J Am Statist Assoc* 84:360–72
- Keiding N, Gill R (1990) Random truncation models and Markov processes. *Ann Statist* 18:582–602
- Lagakos SW, Barraj LM, De Gruttola V (1988) Nonparametric analysis of truncated survival data, with application to AIDS. *Biometrika* 75:515–23
- Rennert L (2018) Statistical methods for truncated survival data. Doctoral dissertation, University of Pennsylvania
- Rennert L, Xie SX (2018) Cox regression model with doubly truncated data. *Biometrics* 74:725–33
- Seaman SR, Presanis A, Jackson C (2022) Estimating a time-to-event distribution from right-truncated data in an epidemic: a review of methods. *Stat Methods Med Res* 31:1641–55
- Shen PS, Liu Y, Maa DP, Ju Y (2017) Analysis of transformation models with right-truncated data. *Statistics* 51:404–18
- Stralkowska-Kominiak E, Stute W (2009) Martingale representations of the Lynden-Bell estimator with applications. *Stat Probabil Lett* 79:814–20
- Sundaram R (2009) Semiparametric inference of proportional odds model based on randomly truncated data. *J Stat Plan Inference* 139:1381–93
- Vakulenko-Lagun B, Mandel M, Betensky RA (2020) Inverse probability weighting methods for Cox regression with right-truncated data. *Biometrics* 76:484–955
- Vakulenko-Lagun B, Mandel M, Betensky RA (2020) Inverse probability weighting methods for cox regression with right-truncated data. *Biometrics* 76:484–95
- Wang M-C, Jewell NP, Tsai W-Y (1986) Asymptotic properties of the product limit estimate under random truncation. *Ann Statist* 14:1597–605
- Wang M-C (1989) A semiparametric model for randomly truncated data. *J Am Statist Assoc* 84:742–8
- Woodroffe M (1985) Estimating a distribution function with truncated data. *Ann Statist* 13:163–77

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.