Quantile residual lifetime analysis for dependent survival and competing risks data

Kwang Woo Ahn

Division of Biostatistics

Medical College of Wisconsin, Milwaukee, WI 53226

email: kwooahn@mcw.edu

Brent R. Logan

Division of Biostatistics

Medical College of Wisconsin, Milwaukee, WI 53226

December 1, 2014

Abstract

The quantile residual lifetime analysis is often performed to evaluate the distributions of remaining lifetimes for survival and competing risks data. The current literature is limited to independent data. We propose a pseudo-value approach to compare quantile residual lifetimes of multiple groups for dependent survival and competing risks data. The pseudo-value approach is extended to dependent event times and dependent censoring times. The empirical Type I errors and statistical power of the proposed study are examined in a simulation study, which shows that the proposed method controls Type I errors very well and has higher power than some existing method. The proposed method is illustrated by a bone marrow transplant data set.

KEYWORDS: Residual lifetime; Pseudo-value; Survival data; Competing risks data

1. INTRODUCTION

Residual life is the residual lifetime of a patient given that the patient survived at least to time t. Statistical inference on residual life may provide patients and clinicians valuable information on evaluating treatments. The quantile residual lifetime is often preferred when the distribution of the residual lifetime is skewed (Ma and Wei 2012). The statistical literature on quantile residual lifetime includes Jung, Jeong and Bandos (2009) and Kim, Zhou and Jeong (2012). Jung et al. (2009) proposed a time-specific log-linear regression model and Kim et al. (2012) studied empirical likelihood inference to test parameters of interest. However, they are restricted to independent survival data.

The cause-specific residual life distribution was proposed by Jeong and Fine (2009) for the competing risks setting. The cause-specific residual life distribution is defined as the residual cumulative incidence function conditional on event-free survival to a given time t (Jeong and Fine 2009). A nonparametric test was developed for testing one sample and two samples (Jeong and Fine 2013). As in residual lifetime analysis for survival data, this is limited to independent data. Statistical inference for comparing multiple groups is also desirable in practice.

The pseudo-value technique has been used for survival and competing risks data (Andersen, Klein and Rosthøj 2003; Logan, Zhang and Klein 2011). Graw, Gerds and Schumacher (2009) further studied the asymptotics of pseudo-value regression for independent data. Ahn and Mendolia (2014) examined comparisons of the median survival distributions using the pseudo-value approach. Relying on generalized estimating equations makes statistical inference on dependent data feasible. Although Logan et al. (2011) studied the pseudo-value technique for dependent event times, it was restricted to independent censoring times. A further study allowing for dependent censoring times needs to be addressed.

We propose a pseudo-value-based method to test the equality of quantile residual lifetimes

of multiple groups for dependent survival and competing risks data. We extend the result of Logan et al. (2011) to dependent censoring in Section 2. In Section 3, we describe the proposed test statistic based on pseudo-values and its asymptotic distribution. A simulation study is performed in Section 4. A bone marrow transplant example is illustrated in Section 5. Finally, we have a brief conclusion in Section 6.

2. PSEUDO-VALUE APPROACH

In this section, we review the pseudo-value approach for competing risks and survival settings and extend it to dependent events and dependent censoring times. First of all, we consider the competing risks setting and define some notations. We assume that there are m clusters and each cluster has ℓ individuals. Let $n = m \times \ell$ be the total sample size. Although the cluster size is fixed at ℓ, as in Spiekerman and Lin (1998) the clusters may have different sizes by defining censoring times as zero when observed times are missing. For simplicity, assume that there are two causes of failure $\epsilon \in \{1, 2\}$. Let T_{ij}, C_{ij}, ϵ_{ij}, and $\mathbf{Z}_{ij}$ be the event time, censoring time, cause of failure, and covariate vector of individual j in cluster i, respectively, for $i = 1, \ldots, m$ and $j = 1, \ldots, \ell$. Let $\mathbf{T}_i = \{T_{ij}, j = 1, \ldots, \ell\}, \mathbf{C}_i = \{C_{ij}, j = 1, \ldots, \ell\}, \boldsymbol{\epsilon}_i = \{\epsilon_{ij}, j = 1, \ldots, \ell\}$, and $\mathbf{Z}_i = \{\mathbf{Z}_{ij}, j = 1, \ldots, \ell\}$. Suppose that $(\mathbf{T}_i, \boldsymbol{\epsilon}_i, \mathbf{C}_i, \mathbf{Z}_i)$ are independent and identically distributed (iid). We assume that the C_{ij}'s do not depend on the $\mathbf{Z}_{ij}$'s and the T_{ij}'s are independent of the C_{ij}'s for $i = 1, \ldots, m$ and $j = 1, \ldots, \ell$. Thus, while event times and censoring times for the same individual are independent, the event times may be correlated within the same cluster. Similarly, the censoring times may be correlated within the same cluster. We further assume that the C_{ij}'s have a common distribution G although censoring times may be correlated within a cluster. Let $X_{ij} = \min(T_{ij}, C_{ij})$ be the observed time.

We consider the marginal cumulative incidence function for cause 1. Let $F_1(t) = P(T \leq t, \epsilon = 1)$ and $N_{kij}(t) = I(T_{ij} \leq t)I(\epsilon_{ij} = k)I(T_{ij} \leq C_{ij})$, where $k = 1, 2$. Define $N_{ij}(t) = N_{1ij}(t) + N_{2ij}(t)$. We further define a risk set indicator $Y_{ij}(t) = I\{t \leq X_{ij}\}$ and $Y(t) = \sum_{i=1}^{m} \sum_{j=1}^{\ell} Y_{ij}$. Following Chen, Kramer, Greene and Rosenberg (2007), we define the em-

pirical cause-specific cumulative hazard functions as $\hat{H}_1(t) = \int_0^t d\hat{H}_1(u)$, where

$$d\hat{H}_1(t) = \sum_{i=1}^{m} \sum_{j=1}^{\ell} \frac{dN_{1ij}(t)}{Y(t)}.$$

Then, the cumulative incidence estimate can be estimated by $\hat{F}_1(t) = \int_0^t \hat{S}(u-) d\hat{H}_1(u)$, where $\hat{S}(u-)$ is the Kaplan-Meier estimate of event-free survival, in which the patient has not experienced either cause 1 or cause 2 (Logan et al. 2011). This estimate is still a consistent estimate of $F_1(t)$ even for dependent competing risks events and dependent censoring times (Zhou and Fine 2012).

A pseudo-value at time t of the jth individual in the ith cluster for $F_1(t)$ is defined by $P_{ij}^f(t) = n\hat{F}_1(t) - (n-1)\hat{F}_1^{-ij}(t)$ for $i = 1, \ldots, m$ and $j = 1, \ldots, \ell$, where $\hat{F}_1^{-ij}(t)$ is the cumulative incidence estimate obtained by omitting the jth individual in the ith cluster. Logan et al. (2011) studied marginal cumulative incidence and survival models for clustered data using the pseudo-value approach. Assuming mutually independent censoring times, they showed that i) $P_{ij}^f(t)$ is approximately independent of $P_{kg}^f(t)$ for $i \neq k$ as $n \to \infty$; and ii) $\lim_{n\to\infty} E(P_{ij}^f(t)|\mathbf{Z}_{ij}) = F_1(t|\mathbf{Z}_{ij})$. The pseudo-values can be used as a response variable in a generalized estimating equation (GEE) setting as described in Andersen et al. (2003), Logan et al. (2011), and Klein and Andersen (2005). Because only a single fixed time point is considered in this paper, we illustrate the use of the GEE at a fixed time point t. To model the marginal cumulative incidence function at time t, we consider $g(F_1(t|\mathbf{Z})) = \boldsymbol{\beta}'\mathbf{Z}$. Let $\boldsymbol{\mu} = F_1(t|\mathbf{Z}) = g^{-1}(\boldsymbol{\beta}'\mathbf{Z})$, $\mathbf{P}_i^f = (P_{i1}^f(t), \ldots, P_{i\ell}^f(t))$, and $\boldsymbol{\mu}_i = (\mu_{i1}(t), \ldots, \mu_{i\ell}(t))$ for $i = 1, \ldots, m$. Then, the GEE is defined as follows:

$$\sum_i \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right)' \mathbf{V}_i^{-1}(\mathbf{P}_i^f - \boldsymbol{\mu}_i) \equiv \sum_i \mathbf{U}_i(\boldsymbol{\beta}) = \mathbf{0},$$

where $\mathbf{V}_i$ is a $\ell \times \ell$ working correlation matrix for cluster i. Then, $\sqrt{m}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ converges in distribution to $N(\mathbf{0}, \Sigma)$, see Logan et al. (2011) and Liang and Zeger (1986). To estimate Σ, the sandwich estimator is used:

$$\hat{\Sigma}_{\boldsymbol{\beta}} = \mathbf{I}(\hat{\boldsymbol{\beta}})^{-1} \widehat{Var}\{\mathbf{U}(\hat{\boldsymbol{\beta}})\} \mathbf{I}(\hat{\boldsymbol{\beta}})^{-1},$$

where

$$\mathbf{I}(\boldsymbol{\beta}) = \sum_i \Big(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\Big)' \mathbf{V}_i^{-1} \Big(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\Big),\ \widehat{Var}\{\mathbf{U}(\hat{\boldsymbol{\beta}})\} = \sum_i \mathbf{U}_i(\hat{\boldsymbol{\beta}})\mathbf{U}_i(\hat{\boldsymbol{\beta}})'.$$

Therefore, dependent competing risks data are readily handled by considering within-cluster correlation between individuals.

Next, we discuss extending this to the setting where the censoring times may also be correlated within a cluster. The cumulative incidence estimate $\hat{F}_1(t)$ can be rewritten as

$$\hat{F}_1(t) = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{\ell} \frac{N_{1ij}(t)}{\hat{G}(X_{ij})},$$

where $\hat{G}(t)$ is the Kaplan-Meier estimate of the censoring survival distribution $G(t)$ obtained by treating censored observations as events (Scheike, Zhang and Gerds 2008). Let $N_{ij}^c(t) = I(C_{ij} \le t)$ and $H^c(t)$ be the cumulative hazard function by treating censored observations as events. Define

$$\pi(t) = \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{\ell} I(X_{ij} \ge t).$$

For dependent event times and dependent censoring times, Zhou and Fine (2012) showed $\hat{G}(t)$ converges in probability to $G(t)$ uniformly on $t \in [0, \mathcal{T}]$ and $\sqrt{m}\{\hat{G}(t) - G(t)\}$ converges weakly to a tight Gaussian process with covariance function $\Sigma^c(s,t) = E\{I_i^c(s)I_i^c(t)\}$, where

$$I_i^c(t) = \sum_{j=1}^{\ell} \int_0^t \frac{1}{\pi(u)} dM_{ij}^c(u),$$

and $M_{ij}^c(t) = N_{ij}^c(t) - \int_0^t I(X_{ij} \ge u) dH^c(u)$. As in Appendix, similarly to Logan et al. (2011), we can show

$$P_{ij}^f(t) = \frac{N_{1ij}(t)}{G(X_{ij})} + \int_0^{X_{ij}} \frac{P(T_f \le t, \epsilon = 1 | T_f \ge u)}{G(u)} dM_{ij}^c(u) + O_p(m^{-1/2}), \tag{1}$$

where T_f is event time of cause 1. Note that the first two terms are the same as those in Equation (2) of Logan et al. (2011). Because censoring times are independent of $\{\mathbf{Z}_{ij}, i = 1, \ldots, m, j = 1, \ldots, \ell\}$, $\lim_{m\to\infty} E\{P_{ij}^f(t)|\mathbf{Z}_{ij}\} = F_1(t|\mathbf{Z}_{ij})$. In addition, $\mathbf{P}_i^f(t) = (P_{i1}^f(t), \ldots, P_{i\ell}^f(t))^T$'s are asymptotically iid for $i = 1, \ldots, m$. The asymptotics of the GEE can be justified similarly to Theorem 2 of Graw et al. (2009). This extends the result of Logan et al. (2011) to dependent censoring times.

For the survival setting, let the survival function $S(t)$ be event-free survival, in which the patient has not experienced any causes. The pseudo-value for survival is defined as $P^s_{ij}(t) = n\hat{S}(t) - (n-1)\hat{S}^{-ij}(t)$ for $i = 1, \ldots, m$ and $j = 1, \ldots, \ell$, where $\hat{S}^{-ij}(t)$ is the Kaplan-Meier estimate obtained by omitting the jth individual in the ith cluster. The consistency of the Kaplan-Meier estimate for dependent events and dependent censoring was shown by Zhou and Fine (2012). Like the competing risks setting, we can show

$$P^s_{ij}(t) = \frac{N_{ij}(t)}{G(X_{ij})} + \int_0^{X_{ij}} \frac{P(T_s \leq t | T_s \geq u)}{G(u)} dM^c_{ij}(u) + O_p(m^{-1/2}), \tag{2}$$

where T_s is event time of any cause. As in the competing risks setting, $\lim_{m\to\infty} E\{P^s_{ij}(t)|\mathbf{Z}_{ij}\} = S(t|\mathbf{Z}_{ij})$ and $\mathbf{P}^s_i(t) = (P^s_{i1}(t), \ldots, P^s_{i\ell}(t))^T$'s are asymptotically iid, which extends the result of Logan et al. (2011) for dependent censoring times. The GEE setting can be justified as shown in Graw et al. (2009).

3. METHOD

In this section, we propose pseudo-value-based methods for testing residual lifetime for competing risks and survival settings and study properties of the proposed methods. Consider the competing risks setting first. Let q_τ be the τth quantile of the cause 1 residual life distribution given event-free survival to t. Jeong and Fine (2009) defined the residual cumulative incidence function given event-free survival to time t for cause 1 as follows:

$$P(T - q_\tau \leq t, \epsilon = 1 | T > t) = \frac{F_1(q_\tau + t) - F_1(t)}{S(t)}.$$

The τth quantile of the cause 1 residual lifetime q_τ given event-free survival to time t satisfies

$$\frac{F_1(q_\tau + t) - F_1(t)}{S(t)} = \tau.$$

Let $A(q_\tau) = F_1(q_\tau + t) - F_1(t) - \tau S(t)$ and $\hat{A}(q_\tau) = \hat{F}_1(q_\tau + t) - \hat{F}_1(t) - \tau\hat{S}(t)$. Jeong and Fine (2009) showed that $A(q_\tau) = 0$ has a unique root. In practice, $\hat{q}_\tau$ is uniquely determined by defining it as the smallest q at which $\{\hat{F}_1(q + t) - \hat{F}_1(t)\}/\hat{S}(t)$ crosses τ (Jeong and Fine 2009), where $\hat{F}_1(t)$ and $\hat{S}(t)$ are the cumulative incidence estimate of cause 1 and the

Kaplan-Meier estimate at time t, respectively. We assume that $F_1(t)$ is absolutely continuous and $f_1(t) = dF_1(t)/dt$ is positive on some neighborhood of $q_\tau + t$. Jeong and Fine (2009) showed the consistency of $\hat{q}_\tau$ for independent data. Similar arguments can be used to show the consistency of $\hat{q}_\tau$ for dependent data as follows: $\hat{A}(q_\tau)$ converges to $A(q_\tau)$ due to the consistency of $\hat{S}(t)$ and $\hat{F}_1(t)$. Because of absolute continuity of $F_1(t)$ and positivity of $f_1(t)$ on some neighborhood of $q_\tau + t$, $A(q_\tau)$ has a unique solution. Thus, $\hat{q}_\tau$ is consistent given τ.

Assume that there are ν groups to compare. Under the null hypothesis, we have $q_{1\tau} = \cdots = q_{\nu\tau} \equiv q_{0\tau}$, where $q_{i\tau}$ is the τth quantile of the cause 1 residual life distribution given event-free survival to t for group i, $i = 1, \ldots, \nu$. Due to the uniqueness of the solution for $A(q_\tau) = 0$, this is equivalent to testing $A(q_{1\tau}) = \cdots = A(q_{\nu\tau}) \equiv A(q_{0\tau}) = 0$.

To compare $A(\cdot)$ values at $q_{0\tau}$ of ν groups, we use the pseudo-value approach. Given q_τ, the pseudo-value for $A(\cdot)$ of individual j in cluster i is defined as $B_{ij}(q_\tau) = \hat{P}^f_{ij}(q_\tau + t) - \hat{P}^f_{ij}(t) - \tau \hat{P}^s_{ij}(t)$ for $i = 1, \ldots, m$ and $j = 1, \ldots, \ell$. Let $q_{0\tau}$ be the solution of $A(x) = 0$. Using (1) and (2), we can show i)

$$E\{B_{ij}(q_{0\tau})|\mathbf{Z}_{ij}\} = F_1(q_{0\tau} + t|\mathbf{Z}_{ij}) - F_1(t|\mathbf{Z}_{ij}) - \tau S(t|\mathbf{Z}_{ij}) + O_p(m^{-1/2}),$$

and ii) $\mathbf{B}_i(q_{0\tau}) = (B_{i1}(q_{0\tau}), \ldots, B_{i\ell}(q_{0\tau}))^T$'s are asymptotically iid for $i = 1, \ldots, m$. The GEE use can be justified as in Theorem 2 of Graw et al. (2009). To apply the GEE, define an indicator variable I_k for group k such that for $k = 1, \ldots, \nu$,

$$I_k = \begin{cases} 1, \text{ if an individual belongs to the } k\text{th group,} \\ 0, \text{ otherwise.} \end{cases}$$

Thus, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_\nu)^T$ is to be estimated. To avoid an identifiability issue, without loss of generality, we fix β_ν at 0 and estimate $\boldsymbol{\beta}_{-\nu} = (\beta_1, \ldots, \beta_{\nu-1})^T$. Let $\hat{q}_{0\tau}$ be the solution of $\hat{A}(q_{0\tau}) = 0$ based on the pooled data. Then, we define pseudo-values as $B_{ij}(\hat{q}_{0\tau})$. Assuming $N_{1ij}(x)$ is continuous at $x = q_{0\tau} + t$ with probability one, under the null hypothesis we can show that $B_{ij}(\hat{q}_{0\tau})$ converges to $B_{ij}(q_{0\tau})$ in probability as in Appendix. An identity link function can be used for the GEE. Testing the null hypothesis $A(q_{1\tau}) = \cdots = A(q_{\nu\tau}) \equiv$

$A(q_{0\tau}) = 0$ is equivalent to testing $\boldsymbol{\beta}_{-\nu} = \mathbf{0}$ given $q_{0\tau}$. Due to the consistency of $\hat{q}_{0\tau}$, the test statistic is given by

$$X^2 = m\hat{\boldsymbol{\beta}}_{-\nu}^{'}\hat{\Sigma}_{\boldsymbol{\beta}_{-\nu}}^{-1}\hat{\boldsymbol{\beta}}_{-\nu},$$

where $\hat{\boldsymbol{\beta}}_{-\nu}$ is found by numerically solving the GEE with $B_{ij}(\hat{q}_{0\tau})$'s and $\hat{\Sigma}_{\boldsymbol{\beta}_{-\nu}}$ is the corresponding sandwich estimate of the covariance matrix of $\hat{\boldsymbol{\beta}}_{-\nu}$. Under the null hypothesis, X^2 follows a chi-squared distribution with degrees of freedom $\nu - 1$.

For the survival setting, let ζ_τ be the τ-quantile residual life function of group i at time t. Then, it satisfies

$$P(T \geq t + \zeta_\tau) = (1-\tau)P(T \geq t) \text{ or } S(t + \zeta_\tau) = (1-\tau)S(t).$$

Define $C(\zeta) = S(t + \zeta_\tau) - (1-\tau)S(t)$. Let $\zeta_{0\tau}$ be the unique solution of $C(\zeta) = 0$. Then, $\hat{\zeta}_{0\tau}$ can be defined as the smallest ζ_τ at which $\hat{S}(t + \zeta_\tau) - (1-\tau)\hat{S}(t)$ crosses zero, where $\hat{S}(t)$ is the Kaplan-Meier estimate at time t based on the pooled data. The consistency of $\hat{\zeta}_{0\tau}$ can be shown similarly to $\hat{q}_{0\tau}$. Assume that there are ν groups to compare. Under the null hypothesis, we have $\zeta_{1\tau} = \cdots = \zeta_{\nu\tau} \equiv \zeta_{0\tau}$, where $\zeta_{i\tau}$ is the τ-quantile residual life function of group i at time t. Like the competing risks setting, this is equivalent to testing $C(\zeta_{1\tau}) = \cdots = C(\zeta_{\nu\tau}) \equiv C(\zeta_{0\tau}) = 0$. The pseudo-value approach can be applied to compare $C(\cdot)$ values at $\zeta_{0\tau}$ of ν groups. Let $D_{ij}(\zeta_\tau) = \hat{P}_{ij}^s(t + \zeta_\tau) - (1-\tau)\hat{P}_{ij}^s(t)$ for $i = 1, \ldots, \ell$ and $j = 1, \ldots, m$. Using (2), we can show i)

$$E\{D_{ij}(\zeta_{0\tau})|\mathbf{Z}_{ij}\} = S(t + \zeta_{0\tau}|\mathbf{Z}_{ij}) - (1-\tau)S(t|\mathbf{Z}_{ij}) + O_p(m^{-1/2}),$$

and ii) $\mathbf{D}_i(\zeta_{0\tau}) = (D_{i1}(\zeta_{0\tau}), \ldots, D_{i\ell}(\zeta_{0\tau}))^T$'s are asymptotically iid for $i = 1, \ldots, m$. Similarly to the proof of convergency of $B_{ij}(\hat{q}_{0\tau})$, it can be shown that $D_{ij}(\hat{\zeta}_{0\tau})$ converges to $D_{ij}(\zeta_{0\tau})$ under the null hypothesis. Using the $D_{ij}(\hat{\zeta}_{0\tau})$'s, we can use the GEE similarly to the competing risks setting to test $C(\zeta_{1\tau}) = \cdots = C(\zeta_{\nu\tau}) \equiv C(\zeta_{0\tau}) = 0$ due to the consistency of $\hat{\zeta}_{0\tau}$.

Table 1: Empirical Type I error rates from comparing four groups for competing risks data

γ	m	Empirical Type I errors
0	100	0.058
	200	0.052
	400	0.051
0.5	100	0.072
	200	0.059
	400	0.056

4. SIMULATION

We perform a simulation study in this section. The simulation size is 5,000 replicates throughout this section. Consider the competing risks setting first. We consider comparing four groups at a significance level $\alpha = 0.05$. A positive stable frailty is used to generate correlated event times and censoring times (Logan et al. 2011). Each cluster has eight individuals, with two individuals in each of the four groups being compared. We consider $m = 100, 200,$ and 400. For each cluster, a random effect w is generated from a positive stable frailty distribution with parameter γ, where the Laplace transformation of the standard positive stable distribution is $L(s) = \exp(-s^\gamma)$. Logan et al. (2011) showed that a marginal model which has proportional subdistribution hazards is obtained by using a positive stable frailty. Two γ values are used: 1 for independent data and 0.5 for dependent data. As discussed in Logan et al. (2011), dependent competing risks data for each cluster can be obtained from

$$F_1(t|w) = 1 - \{1 - p(1 - e^{-t})\}^w,$$
$$F_2(t|w) = (1 - p)^w(1 - e^{-t}).$$

Table 2: Comparison to Jeong and Fine (2013) for competing risks data with two groups

		Empirical Type I error		Empirical power	
γ	m	JF	PM	JF	PM
0	50	0.020	0.052	0.095	0.174
	100	0.014	0.050	0.155	0.308
	200	0.018	0.050	0.362	0.524
0.5	50	0.034	0.057	0.056	0.215
	100	0.008	0.050	0.110	0.363
	200	0.004	0.050	0.317	0.634

Independent of event times, the corresponding censoring times for each cluster are generated from

$$G(t|w_c) = e^{-\lambda w_c t},$$

where a random effect w_c is generated from a positive stable frailty distribution with two γ values: 1 for independent data and 0.5 for dependent data. We set p to 0.5 and 0.27 for $\gamma = 1$ and 0.5, respectively. We select λ to generate 25% of events with cause 1, 25% of events with cause 2, and 50% of censoring. Given a patient survived event free to at least time $t = 0.3$, the 0.25th quantile of the cause 1 residual lifetimes are compared between the four groups. The true 0.25th quantile of the cause 1 residual lifetimes are $\log(2)$ for $\gamma = 1$ and 1.327 for $\gamma = 0.5$ in all four groups. The identity link function with an independence working correlation matrix is used for the pseudo-value approach. We also examined the exchangeable working correlation matrix and the unstructured working correlation matrix, but there was negligible difference from the result with the independence working correlation matrix. Table 1 shows the summary of 5,000 iterations. It shows that the proposed method controls Type I error rates very well. As m increases, the empirical Type I error rates become closer to 0.05.

Next, we compare the performance of the proposed method and Jeong and Fine (2013). To compare Type I error rates, we use the same setting as above except that we consider comparing two groups instead of four groups because Jeong and Fine (2013) is restricted to testing two groups. As a result, each cluster has four individuals, with two individuals in each of the two groups being compared. Table 2 summarizes the results. "JF" and "PM" indicate Jeong and Fine (2013) and the proposed method, respectively. The proposed method controls Type I error rates very well for independent and dependent data. Jeong and Fine (2013) is somewhat conservative, which was also observed in the simulation studies of Jeong and Fine (2013). It appears that Jeong and Fine (2013) becomes more conservative for dependent data under our simulation setting. To examine statistical power, we generate event times with i) when $\gamma = 1$, $p = 0.4$ for group 1 and $p = 0.6$ for group 2; and ii) when $\gamma = 0.5$, $p = 0.3$ for group 1 and $p = 0.8$ for group 2. We choose λ to generate 50% of censoring. The true 0.25th quantiles of the cause 1 residual lifetimes for groups 1 and 2 are $\log(8/3)$ and $\log(12/7)$ for $\gamma = 1$, and 1.152 and 0.429 for $\gamma = 0.5$. As we can see from Table 2, the pseudo-value approach has higher power than Jeong and Fine (2013). Compared to the independent data, the difference of empirical statistical powers between the proposed test and Jeong and Fine (2013) is larger in the dependent data. This is likely because the pseudo-value-based test effectively utilizes the within cluster correlation in this stratified simulation design.

For the survival setting, we consider four distributions for survival and censoring distributions: exponential distribution, Weibull distribution, Gompertz distribution, and log-logistic distribution. For $x \geq 0$, their probability density functions are

- Exponential distribution: $\lambda_1 \exp(-\lambda_1 x)$;
- Weibull distribution: $2\lambda_2 x \exp(-\lambda_2 x^2)$;
- Gompertz distribution: $\lambda_3 \exp(x/3) \exp[3\lambda_3\{1 - \exp(x/3)\}]$;
- Log-logistic distribution: $\lambda_4/(1 + \lambda_4 x)^2$.

The corresponding survival functions are $\exp(-\lambda_1 x), \exp(-\lambda_2 x^2), \exp[3\lambda_3\{1-\exp(x/3)\}]$, and $1/(1+\lambda_4 x)$, respectively. We compare four groups to examine empirical Type I error rates at the significance level $\alpha = 0.05$. Each cluster is assumed to have eight individuals with two individuals in each of the four groups being compared. We consider $m = 100, 200,$ and 400. The identity link function with an independence working correlation matrix is used for the pseudo-value approach as in the competing risks setting. The exchangeable working correlation matrix and the unstructured working correlation matrix were also examined, but there was negligible difference from the result with the independence working correlation matrix as in the competing risks setting.

Normal copulas are employed to generate correlated survival times and censoring times within each cluster. The 8×8 exchangeable correlation matrix $\mathcal{C}$ with correlation $\rho = 0$ and 0.5 is used for the normal copulas, i.e.,

$$\mathcal{C} = \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{pmatrix}.$$

Thus, $\rho = 0$ means that the survival and censoring times of the four groups are mutually independent. On the other hand, the survival and censoring times within the same cluster are correlated with $\rho = 0.5$. Using eight-dimensional random vectors on the unit cube $[0,1]^8$ from normal copulas given ρ, the survival times are generated corresponding to their marginal survival distributions. Independent of the survival times, the censoring times are generated using normal copulas with the same ρ that is used for survival times. For the detailed use of copulas, see Yan (2007).

Table 3: Empirical Type I error rates when the survival distributions of the four groups are equal.

ρ	m	EEEE	GGGG	LLLL	WWWW
0	100	0.056	0.060	0.053	0.056
	200	0.052	0.054	0.054	0.051
	400	0.050	0.051	0.055	0.049
0.5	100	0.061	0.066	0.063	0.059
	200	0.062	0.062	0.058	0.053
	400	0.051	0.053	0.052	0.051

Table 4: Empirical Type I error rates when the survival distributions of the two groups are different from the other two groups' survival distributions.

ρ	m	EEGG	EELL	EEWW	GGLL	GGWW	WWLL
0	100	0.058	0.053	0.052	0.053	0.062	0.046
	200	0.053	0.052	0.054	0.048	0.056	0.046
	400	0.053	0.053	0.051	0.047	0.052	0.046
0.5	100	0.067	0.054	0.057	0.059	0.062	0.051
	200	0.060	0.057	0.053	0.057	0.055	0.052
	400	0.052	0.049	0.048	0.052	0.053	0.048

Table 5: Comparison to Jeong and Fine (2013) for survival data with two groups

		Empirical Type I error						Empirical power					
		EE		GG		EG		EE		GG		EG	
ρ	m	JF	PM	JF	PM	JF	PM	JF	PM	JF	PM	JF	PM
0	50	0.019	0.058	0.021	0.057	0.021	0.055	0.094	0.164	0.179	0.273	0.123	0.211
	100	0.017	0.046	0.019	0.046	0.017	0.048	0.171	0.268	0.352	0.458	0.224	0.350
	200	0.021	0.050	0.021	0.050	0.021	0.051	0.343	0.454	0.662	0.746	0.466	0.590
0.5	50	0.010	0.061	0.014	0.047	0.012	0.062	0.074	0.182	0.156	0.303	0.102	0.234
	100	0.010	0.050	0.010	0.055	0.011	0.052	0.154	0.299	0.339	0.514	0.211	0.399
	200	0.011	0.052	0.014	0.047	0.012	0.048	0.337	0.520	0.678	0.807	0.461	0.659

To examine empirical Type I errors at the significance level $\alpha = 0.05$, survival times are generated from exponential distribution, Weibull distribution, Gompertz distribution, and log-logistic distribution with $\lambda_1 = 2/3 \times \log 2, \lambda_2 = 4/15 \times \log 2, \lambda_3 = (\log 2)/3/\{\exp(2/3) - \exp(1/6)\}$, and $\lambda_4 = 1$, respectively. The corresponding censoring times are generated from the same distribution that is used for survival times, which leads to 50% of censoring rate. Given that a patient survived event free at least to time $t = 0.5$, the true residual survival median $\zeta_{0\tau}$ is 1.5 for each survival distribution, where $\tau = 0.5$. Survival probabilities at $t = 0.5$ are $\exp\{-(\log 2)/3\}, \exp\{-(\log 2)/15\}$,
$\exp[(\log 2)\{1 - \exp(6)\}/\{\exp(2/3) - \exp(1/6)\}]$, and 2/3 for the exponential distribution, Weibull distribution, Gompertz distribution, and log-logistic distribution, respectively. The residual survival median of the four groups are compared given a patient survived event free at least to time $t = 0.5$. Tables 3 and 4 show empirical Type I error rates i) assuming four groups have the same distributions; and ii) assuming two groups have a different survival distribution from the other two groups. E, W, G, and L indicate the exponential distribution, Weibull distribution, Gompertz distribution, and log-logistic distribution, respectively. For example, EEGG indicates that two groups have the exponential survival distributions and

the other two groups have the Gompertz distributions. The proposed method controls Type I error rates very well for independent and dependent survival data. As the number of clusters increases, the empirical Type I error rates become closer to 0.05 in general.

To compare the proposed method to Jeong and Fine (2013), we consider two-group comparison. The two-sample test statistic of Jeong and Fine (2013) for competing risks data is modified by assuming that there are no competing risks. We examine three cases: i) two groups have exponential survival distributions (EE); ii) two groups have Gompertz survival distributions (GG); and iii) one group has an exponential survival distribution and the other group has a Gompertz survival distribution (EG). Censoring times are generated from the same distribution that is used for survival times to generate 50% of censoring rate. To evaluate Type I error rates, we use the same parameters for each survival distribution as those in the Type I error testing of the four-group comparison. Table 5 shows the summary of 5,000 replicates. While the method of Jeong and Fine (2013) tends to be conservative, the proposed method controls Type I errors very well. As the number of clusters increases, the empirical Type I error rates of the proposed method get closer to 0.05 in general. It appears that the empirical Type I error rates of Jeong and Fine (2013) for $\rho = 0.5$ become smaller than those for $\rho = 0$ because of dependent event and censoring times. To compare statistical power between Jeong and Fine (2013) and the proposed method, we consider that the median residual lifetimes of the two groups given that a patient survived event free at least to time $t = 0.5$ are 1.5 and 2, respectively. The parameters for each survival distribution are chosen accordingly. Censoring times are generated from the same distribution that is used for survival times, which results in 50% of censoring rate. As can be seen in Table 5, the proposed method has higher statistical power than Jeong and Fine (2013).

5. EXAMPLE

The data for illustration were collected by the Center for International Blood and Marrow Transplant Research (Shaw, Kan, Spellman, Aljurf, Ayas, Burke, Cairo, Chen, S. M. Davies, Gajewski, Gale, Godder, Hale, Heemskerk, Horan, Kamani, Kasow, Chan, Lee, Leung,

Lewis, Miklos, Oudshoorn, Petersdorf, Ringden, Sanders, Schultz, Seber, Setterholm, Wall, Yu and Pulsipher 2010) and consisted of pediatric patients with myelodysplastic syndrome undergoing a first allogeneic transplantation from 1993 to 2006. The study population in this example consisted of 1453 patients from 105 transplant centers. Three disease groups were compared for analysis: acute lymphoblastic leukemia (ALL), acute myelogenous leukemia (AML), and chronic myelogenous leukemia (CML). There were 553 patients with AML, 756 patients with ALL, and 144 patients with CML. Disease-free survival (DFS) and relapse were outcomes of interest in this analysis. Treatment-related mortality was treated as the competing risk for relapse. Significant center effects on DFS and relapse rates were found using the random effect score test of Commenges and Andersen (1995). Their p-values were smaller than 0.001. Censoring times for LFS and relapse were also correlated because their p-values from the score test by treating censoring observations as events were less than 0.001. The independence working correlation matrix with the identity link function was used for the GEE with the pseudo-values.

First of all, the 0.25th quantile residual lifetimes of LFS given patients survived disease free at least one year were compared between the three diseases. The two upper plots of Figure 1 show the Kaplan-Meier DFS curves and cumulative incidence curves of relapse of the three disease groups. The lower-left plot shows the estimated residual disease-free survival (RDFS) of the three groups given disease-free survival to at least one year, where $\text{RDFS}(t) = S(12+t)/S(12)$. The dotted horizontal line indicates $\text{RDFS} = (1-\tau) = 0.75$, where $\tau = 0.25$. The estimated 0.25th quantile residual lifetimes of AML, ALL, and CML were 121, 71, and 18 months, respectively. The 0.25th quantile residual lifetime of CML appears to be different from AML and ALL. The proposed pseudo-value approach found a significant difference with p-value 0.038 at the significance level 0.05. Next, we considered 0.25th quantile cause-specific residual lifetimes of relapse given patients survived disease free to at least six months. The lower-right plot of Figure 1 shows the residual cumulative

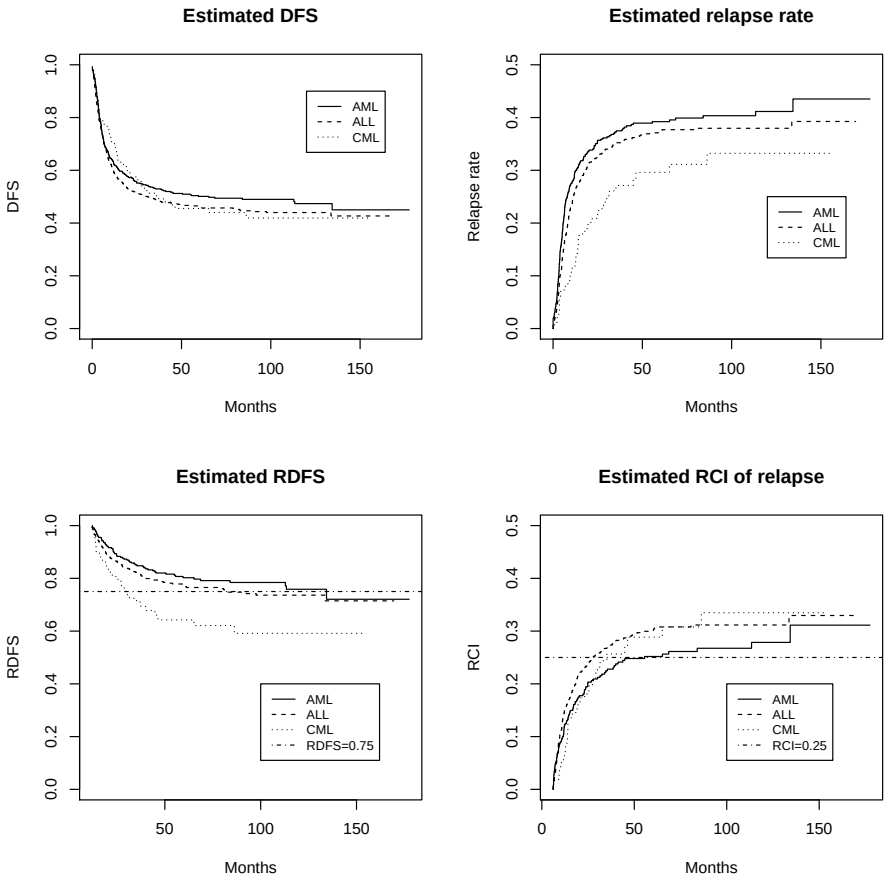

Figure 1: Estimated DFS and relapse rates for disease groups

incidence (RCI) curves of relapse given disease-free survival to at least six months for the three disease groups, where $\text{RCI}(t) = \{F_1(t+6) - F_1(6)\}/S(6)$. The dotted horizontal line represents RCI = 0.25. The estimated 0.25th quantile cause-specific residual lifetimes of AML, ALL, and CML were 49, 22, and 29 months, respectively. The p-value from the proposed method was 0.314, which was not statistically significant.

6. CONCLUSION

We have proposed the pseudo-value approach to compare residual lifetimes for survival and competing risks data. The pseudo-value approach was extended to dependent event times and dependent censoring times assuming that event times are independent of censoring times and censoring times do not depend on any covariates. The simulation study showed that the proposed method controlled Type I errors satisfactorily for independent and dependent data and had higher power than Jeong and Fine (2013). A bone marrow transplant data set was illustrated as an example. An interesting future research problem includes a pseudo-value approach for residual life to adjust for multiple variables.

REFERENCES

Ahn, K. W., and Mendolia, F. (2014), "Pseudo–value approach for comparing survival medians for dependent data," *Statistics in Medicine*, 33, 1531–1538.

Andersen, P. K., Klein, J. P., and Rosthøj, S. (2003), "Generalised linear models for correlated pseudo-observations, with applications to multi-state models," *Biometrika*, 90, 15–27.

Chen, B. E., Kramer, J. L., Greene, M. H., and Rosenberg, P. S. (2007), "Competing risks analysis of correlated failure time data," *Biometrics*, 64, 172–179.

Commenges, D., and Andersen, P. K. (1995), "Score test of homogeneity for survival data," *Lifetime Data Analysis*, 1, 145–156.

Graw, F., Gerds, T. A., and Schumacher, M. (2009), "On pseudo-values for regression analysis in competing risks models," *Lifetime Data Analysis*, 15, 241–255.

Jeong, J. H., and Fine, J. (2009), "A note on cause-specific residual life," *Biometrika*, 96, 237–242.

Jeong, J. H., and Fine, J. (2013), "Nonparametric inference on cause-specific quantile residual life.," *Biometrical Journal*, 55, 68–81.

Jung, S. H., Jeong, J. H., and Bandos, H. (2009), "Regression on Quantile Residual Life," *Biometrics*, 65, 1203–1212.

Kim, M. O., Zhou, M., and Jeong, J. H. (2012), "Censored quantile regression for residual lifetimes," *Lifetime Data Analysis*, 18, 177–194.

Klein, J. P., and Andersen, P. K. (2005), "Regression modeling of competing risks data based on pseudo-values of the cumulative incidence function," *Biometrics*, 61, 223–229.

Liang, K. Y., and Zeger, S. L. (1986), "Longitudinal data analysis using generalized linear models," *Biometrika*, 73, 13–22.

Logan, B., Zhang, M. J., and Klein, J. P. (2011), "Marginal models for clustered time to event data with competing risks using pseudovalues," *Biometrics*, 67, 1–7.

Ma, Y., and Wei, Y. (2012), "Analysis on censored quantile residual life model via spline smoothing," *Statistica Sinica*, 22, 47–68.

Scheike, T. H., Zhang, M. J., and Gerds, T. A. (2008), "Predicting cumulative incidence probability by direct binomial regression," *Biometrika*, 95, 205–220.

Shaw, P. J., Kan, F., Spellman, K. W. A. S. R., Aljurf, M., Ayas, M., Burke, M., Cairo, M. S., Chen, A. R., S. M. Davies, H. F., Gajewski, J., Gale, R. P., Godder, K., Hale, G. A., Heemskerk, M. B., Horan, J., Kamani, N., Kasow, K. A., Chan, K. W., Lee, S. J., Leung, W. H., Lewis, V. A., Miklos, D., Oudshoorn, M., Petersdorf, E. W.,

Ringden, O., Sanders, J., Schultz, K. R., Seber, A., Setterholm, M., Wall, D. A., Yu, L., and Pulsipher, M. A. (2010), "Outcomes of pediatric bone marrow transplantation for leukemia and myelodysplasia using matched sibling, mismatched related, or matched unrelated donors," *Blood*, 116, 4007–4015.

Spiekerman, C. F., and Lin, D. Y. (1998), "Marginal regression models for multivariate failure time data," *Journal of the American Statistical Association*, 93, 1164–1175.

Yan, J. (2007), "Enjoy the joy of copulas: with a package copula," *Journal of Statistical Software*, 21, 1–21.

Zhou, B., and Fine, J. (2012), "Competing risks regression for clustered data," *Biostatistics*, 13, 371–383.

ACKNOWLEDGMENTS

The US National Science Foundation (DMS-1021896) and the US National Cancer Institute (U24CA076518) partially supported this work. The authors would like to thank Ms. Elizabeth Cho for her proof reading and Dr. Jong-Hyeon Jeong for sharing his R codes of Jeong and Fine (2013).

APPENDIX

A. Proof of (1) and (2)

We prove (1) and (2) by following the arguments of Logan et al. (2011). We consider the competing risks setting to show (1). The proof of (2) can be similarly done. We have

$$\begin{aligned}
P_{ij}^{f}(t) &= n\hat{F}_1(t) - (n-1)\hat{F}_1^{-ij}(t) \\
&= \sum_{a=1}^{m}\sum_{b=1}^{\ell}\frac{N_{1ab}(t)}{\hat{G}(X_{ab})} - \sum_{(a,b)\neq(i,j)}\sum\frac{N_{1ab}(t)}{\hat{G}^{-ij}(X_{ab})} \\
&= \frac{N_{1ij}(t)}{G(X_{ij})} + \sum_{(a,b)\neq(i,j)}\sum N_{1ab}(t)\Big[\Big\{\frac{1}{\hat{G}(X_{ab})} - \frac{1}{G(X_{ab})}\Big\} - \Big\{\frac{1}{\hat{G}^{-ij}(X_{ab})} - \frac{1}{G(X_{ab})}\Big\}\Big] \\
&\quad + N_{1ij}(t)\Big\{\frac{1}{\hat{G}(X_{ij})} - \frac{1}{G(X_{ij})}\Big\}.
\end{aligned}$$

The third term $N_{1ij}(t)\{1/\hat{G}(X_{ij}) - 1/G(X_{ij})\}$ is $O_p(m^{-1/2})$ due to the consistency of $\hat{G}(t)$. Consider the second term. Let

$$R(t) = \sum_{a=1}^{m}\sum_{b=1}^{\ell} I(X_{ab} \geq t),\ M^c(t) = \sum_{a=1}^{m}\sum_{b=1}^{\ell} M^c_{ab}(t),\ M^{c(-ij)}(t) = \sum_{(a,b)\neq(i,j)}\sum M^c_{ab}(t).$$

Using

$$\frac{1}{\hat{G}(X_{ab})} - \frac{1}{G(X_{ab})} = \frac{1}{\hat{G}(X_{ab})}\int_0^{X_{ab}} \frac{\hat{G}(u-)}{G(u)R(u)} dM^c(u),$$

the second term is equal to

$$\sum_{(a,b)\neq(i,j)}\sum N_{1ab}(t)\Big\{\frac{1}{\hat{G}(X_{ab})}\int_0^{X_{ab}} \frac{\hat{G}(u-)}{G(u)R(u)} dM^c(u) - \frac{1}{\hat{G}^{-ij}(X_{ab})}\int_0^{X_{ab}} \frac{\hat{G}^{-ij}(u-)}{G(u)R(u)} dM^{c(-ij)}(u)\Big\}.$$

Thus, the second term becomes

$$\begin{aligned}
&\sum_{(a,b)\neq(i,j)}\sum \frac{N_{1ab}(t)}{\hat{G}(X_{ab})\hat{G}^{-ij}(X_{ab})} \\
&\quad\times \int_0^{X_{ab}} \Big\{\frac{\hat{G}^{-ij}(X_{ab})\hat{G}(u-)}{G(u)R(u)} - \frac{\hat{G}(X_{ab})\hat{G}^{-ij}(u-)}{G(u)R^{-ij}(u)}\Big\} dM^{c(-ij)}(u) \qquad (3)\\
&\quad+ \sum_{(a,b)\neq(i,j)}\sum \frac{N_{1ab}(t)}{\hat{G}(X_{ab})}\int_0^{X_{ab}} \frac{\hat{G}(u-)}{G(u)R(u)} dM^c_{ij}(u).
\end{aligned}$$

Consider the first term of (3). It can be seen that

$$\begin{aligned}
&\frac{n(n-1)}{\sqrt{m}}\int_0^{X_{ab}} \Big\{\frac{\hat{G}^{-ij}(X_{ab})\hat{G}(u-)}{G(u)R(u)} - \frac{\hat{G}(X_{ab})\hat{G}^{-ij}(u-)}{G(u)R^{-ij}(u)}\Big\} dM^{c(-ij)}(u) \\
&= \int_0^{X_{ab}} \Big\{\frac{(n-1)\hat{G}^{-ij}(X_{ab})\hat{G}(u-)}{G(u)R(u)/n} - \frac{n\hat{G}(X_{ab})\hat{G}^{-ij}(u-)}{G(u)R^{-ij}(u)/(n-1)}\Big\} \frac{dM^{c(-ij)}(u)}{\sqrt{m}}
\end{aligned}$$

converges in distribution to some random variable K_{ab}. Because $n = m \times \ell$, the first term of (3) is asymptotically equivalent to

$$\frac{\sqrt{m}}{n}\sum_{(a,b)\neq(i,j)}\sum \frac{N_{1ab}(t)K_{ab}(t)}{(n-1)\hat{G}(X_{ab})\hat{G}^{-ij}(X_{ab})} = O_p(m^{-1/2}).$$

We can show that the second term of (3) is asymptotically equivalent to

$$\sum_{(a,b)\neq(i,j)}\sum \frac{N_{1ab}(t)}{G(X_{ab})}\Big[\int_0^{X_{ab}} \frac{1}{R(u)} dM^c_{ij}(u)\Big].$$

Using the definition of $M_{ij}^c(t)$, we have

$$\int_0^{X_{ab}} \frac{1}{R(u)} dM_{ij}^c(u) = \int_0^{X_{ij}} \frac{I(u \le X_{ab})}{R(u)} dM_{ij}^c(u).$$

Thus, we have

$$\sum_{(a,b)\neq(i,j)} \sum \frac{N_{1ab}(t)}{G(X_{ab})} \Big[\int_0^{X_{ab}} \frac{1}{R(u)} dM_{ij}^c(u) \Big] = \int_0^{X_{ij}} \frac{1}{R(u)/n} \frac{1}{n} \sum_{(a,b)\neq(i,j)} \sum \frac{N_{1ab}(t) I(X_{ab} \ge u)}{G(X_{ab})} dM_{ij}^c(u).$$

Consider

$$\frac{1}{n} \sum_{(a,b)\neq(i,j)} \sum \frac{N_{1ab}(t) I(X_{ab} \ge u)}{G(X_{ab})}. \tag{4}$$

By the law of large numbers, (4) converges in probability to

$$E\Big\{ \frac{N_{1ab}(t) I(X_{ab} \ge u)}{G(X_{ab})} \Big\}.$$

We have

$$\begin{aligned}
E\Big\{ \frac{N_{1ab}(t) I(X_{ab} \ge u)}{G(X_{ab})} \Big\} &= E\Big[E\Big\{ \frac{I(T_{ab} \le t) I(\epsilon_{ab} = 1) I(T_{ab} \le C_{ab}) I(X_{ab} \ge u)}{G(X_{ab})} \Big| T_{ab} \Big\} \Big] \\
&= E\Big[E\Big\{ \frac{I(T_{ab} \le t) I(\epsilon_{ab} = 1) I(T_{ab} \le C_{ab}) I(T_{ab} \ge u)}{G(T_{ab})} \Big| T_{ab} \Big\} \Big] \\
&= E\Big[\frac{I(T_{ab} \le t) I(\epsilon_{ab} = 1) I(T_{ab} \ge u)}{G(T_{ab})} E\Big\{ I(T_{ab} \le C_{ab}) \Big| T_{ab} \Big\} \Big] \\
&= E\{ I(T_{ab} \le t) I(\epsilon_{ab} = 1) I(T_{ab} \ge u) \} \\
&= P(u \le T_f \le t, \epsilon = 1).
\end{aligned}$$

Note that $R(t)/n$ converges to $P(T_f \ge u)P(C \ge u) = P(T_f \ge u)G(u)$. Then, the third term is asymptotically equivalent to

$$\int_0^{X_{ij}} \frac{P(T_f \le t, \epsilon = 1 | T_f \ge u)}{G(u)} dM_{ij}^c(u),$$

which completes the proof of (1).

B. Proof of convergence of $B_{ij}(\hat{q}_{0\tau})$

Because $B_{ij}(\hat{q}_{0\tau}) = \hat{P}_{ij}^f(\hat{q}_\tau + t) - \hat{P}_{ij}^f(t) - \tau \hat{P}_{ij}^s(t)$, it is sufficient to show that $\hat{P}_{ij}^f(\hat{q}_\tau + t)$ converges to $\hat{P}_{ij}^f(q_\tau + t)$. Using (1), $P_{ij}^f(\hat{q}_\tau + t)$ is asymptotically equivalent to

$$\frac{N_{1ij}(\hat{q}_\tau + t)}{G(X_{ij})} + \int_0^{X_{ij}} \frac{P(T_f \le \hat{q}_\tau + t, \epsilon = 1 | T_f \ge u)}{G(u)} dM_{ij}^c(u).$$

Because $N_{1ij}(x)$ is continuous at $x = q_\tau + t$ with probability one. Therefore, noting that $\hat{q}_\tau$ is consistent, the first term converges in probability to

$$\frac{N_{1ij}(q_\tau + t)}{G(X_{ij})}.$$

Because of the continuity of $F_1(t)$ and consistency of $\hat{q}_\tau$, the second term converges to

$$\int_0^{X_{ij}} \frac{P(T_f \leq q_\tau + t, \epsilon = 1 | T_f \geq u)}{G(u)} dM_{ij}^c(u),$$

which completes the proof.