

APPENDIX A. Supplementary Material

APPENDIX A.1. Proofs for the theorems in the manuscript

Proof for Theorem 1. Let $M_{jk}(t) = N_{jk}(t) - \int_0^t Y_k(u) d\Lambda_{jk}(u)$, where $\Lambda_{jk}(t) = \int_0^t \lambda_{jk}(u) du$ is the cumulative cause-specific hazard for cause j in group k . Under the null hypothesis, we can rewrite (3) as

$$n^{-1/2}U_{11} = \int_0^\tau W_1(t) \frac{Y_1(t)Y_2(t)}{Y_{\cdot}(t)} \left\{ \frac{dM_{11}(t)}{Y_1(t)} - \frac{dM_{12}(t)}{Y_2(t)} \right\},$$

and (6) as

$$n^{-1/2}\tilde{U}_{11} = \sum_{j=1}^2 \sum_{k=1}^2 \{A_{jk}(\tau) + C_k(\tau)B_{jk}(\tau)\},$$

where

$$\begin{aligned} A_{jk}(\tau) &= \int_0^\tau [D_{jk}(t) - E_{jk}(t)C_k(t)] \hat{h}_k^{-1}(t) n^{-1/2} dM_{jk}(t), \\ B_{jk}(\tau) &= \int_0^\tau E_{jk}(t) \hat{h}_k^{-1}(t) n^{-1/2} dM_{jk}(t), \\ C_k(\tau) &= \int_0^\tau n^{-1} \tilde{W}(t) R_1(t) [I(k=1) - R_k(t)/R_{\cdot}(t)] / \hat{G}_{1k}(t-) dF_{1k}(t), \\ D_{jk}(\tau) &= I(j=1) n^{-1} \tilde{W}(\tau) R_1(\tau) [I(k=1) - R_k(\tau)/R_{\cdot}(\tau)] / \hat{G}_{1k}(\tau-), \\ E_{jk}(\tau) &= I(j=1) - G_{1k}(\tau)/S_k(\tau). \end{aligned}$$

Under some regularity conditions, by using the multivariate martingale central limiting theorem (Fleming and Harrington (1991), Theorem 5.3.5), we can prove that $n^{-1/2}(U_{11}, \tilde{U}_{11})^T$ has a multivariate normal limiting distribution with mean $\mathbf{0}$ and variance-covariance $\Sigma^{(1)} = (\sigma_{ij}^{(1)})$, where $\sigma_{11}^{(1)}$ and $\sigma_{22}^{(1)}$ are developed by Fleming and Harrington (1991); Gray (1988), where

$$\begin{aligned} \sigma_{11}^{(1)} &= \sigma^2 = \int_0^\tau w_1^2(t) \frac{y_1(t)y_2(t)}{y_{\cdot}(t)} d\Lambda_{11}(t), \\ \sigma_{22}^{(1)} &= \tilde{\sigma}^2 = \sum_{k=1}^2 n^{-1} \left\{ \int_0^{\tau_1} a_k^2(t) h_k^{-1}(t) h_{\cdot}^{-1}(t) dF_{1\cdot}(t) + \int_0^{\tau_1} b_{2k}^2(t) h_k^{-2}(t) dF_{2k}(t) \right\}, \end{aligned} \tag{A.1}$$

with

$$\begin{aligned}
\Lambda_{jk}(t) &= \int_0^t \lambda_{jk}(u) du, \\
a_k(t) &= d_{jk}(t) + b_{jk}(t), \\
b_{jk}(t) &= [I(j=1) - G_{1k}(t)/S_k(t)] [c_k(\tau_1) - c_k(t)], \\
c_k(t) &= \int_0^t d_{1k}(u) \tilde{\lambda}_{1k}(u) du, \\
d_{jk}(t) &= I(j=1) \tilde{W}(t) R_1(t) [I(k=1) - h_k(t)/h_*(t)] / G_{1k}(t), \\
h_k(t) &= I(t \leq \tau_k) y_k(t) / S_k(t), \\
h_*(t) &= I(t \leq \max(\tau_1, \tau_2)) (y_1(t) + y_2(t)) / S_k(t), \\
y_k(t) &= p_k S_k(t) S_k^c(t), \\
p_k &= n_k / (n_1 + n_2).
\end{aligned}$$

To obtain the covariance $\sigma_{12}^{(1)}$, we first note that

$$\begin{aligned}
& \left\langle n^{-1/2} U_{11}, n^{-1/2} \tilde{U}_{11} \right\rangle \\
&= n^{-1} \left\langle \int_0^\tau W_1(t) \frac{Y_1(t) Y_2(t)}{Y_*(t)} \left\{ \frac{dM_{11}(t)}{Y_1(t)} - \frac{dM_{12}(t)}{Y_2(t)} \right\}, \sum_{k=1}^2 \sum_{j=1}^2 \{A_{jk}(\tau) + C_k(\tau) B_{jk}(\tau)\} \right\rangle \\
&= n^{-1} \left\langle \int_0^\tau W_1(t) \frac{Y_1(t) Y_2(t)}{Y_*(t)} \left\{ \frac{dM_{11}(t)}{Y_1(t)} - \frac{dM_{12}(t)}{Y_2(t)} \right\}, \right. \\
&\quad \left. \int_0^t V_{11}(t) dM_{11}(t) + C_1(\tau) \int_0^\tau E_{11}(t) \hat{h}_1^{-1}(t) dM_{11}(t) \right. \\
&\quad \left. + \int_0^\tau V_{12}(t) dM_{12}(t) + C_2(\tau) \int_0^\tau E_{12}(t) \hat{h}_2^{-1}(t) dM_{12}(t) \right\rangle \\
&= n^{-1} \left\{ \int_0^\tau W_1(t) \frac{Y_2(t)}{Y_*(t)} V_{11}(t) + C_1(\tau) \int_0^\tau W_1(t) \frac{Y_2(t)}{Y_*(t)} E_{11}(t) \hat{h}_1^{-1}(t) \right\} d\langle M_{11}, M_{11} \rangle(t) \\
&\quad + n^{-1} \left\{ \int_0^\tau W_1(t) \frac{Y_1(t)}{Y_*(t)} V_{12}(t) + C_2(\tau) \int_0^\tau W_1(t) \frac{Y_1(t)}{Y_*(t)} E_{12}(t) \hat{h}_2^{-1}(t) \right\} d\langle M_{12}, M_{12} \rangle(t),
\end{aligned} \tag{A.2}$$

where $V_{jk}(t) = [D_{jk}(t) - E_{jk}(t) C_k(t)] \hat{h}_k^{-1}(t)$. Furthermore, $M_{jk}(t)$ are orthogonal square integrable martingales with predictable variation process

$$\langle M_{jk}(t), M_{j'k'}(t) \rangle = \gamma_{jj'} \gamma_{kk'} \int_0^t Y_k(u) d\Lambda_{jk}(u), \tag{A.3}$$

where $\gamma_{uv} = 1$ if $u = v$. After plugging (A.3) into (A.2), we have

$$\begin{aligned} & \left\langle n^{-1/2}U_{11}, n^{-1/2}\tilde{U}_{11} \right\rangle \\ &= n^{-1} \left[\int_0^\tau W_1(t) \frac{Y_2(t)}{Y(t)} V_{11}(t) + C_1(\tau) \int_0^\tau W_1(t) \frac{Y_2(t)}{Y(t)} E_{11}(t) \hat{h}_1^{-1}(t) \right] Y_1(t) d\Lambda_{11}(t) \\ & \quad + n^{-1} \left[\int_0^\tau W_1(t) \frac{Y_1(t)}{Y(t)} V_{12}(t) + C_2(\tau) \int_0^\tau W_1(t) \frac{Y_1(t)}{Y(t)} E_{12}(t) \hat{h}_2^{-1}(t) \right] Y_2(t) d\Lambda_{12}(t), \end{aligned}$$

which converges in probability to

$$\begin{aligned} \sigma_{12}^{(1)} &= \left[\int_0^\tau w_1(t) \frac{y_2(t)}{y(t)} v_{11}(t) + c_1(\tau) \int_0^\tau w_1(t) \frac{y_2(t)}{y(t)} e_{11}(t) h_1^{-1}(t) \right] y_1(t) d\Lambda_{11}(t) \\ & \quad + \left[\int_0^\tau w_1(t) \frac{y_1(t)}{y(t)} v_{12}(t) + c_2(\tau) \int_0^\tau w_1(t) \frac{y_1(t)}{y(t)} e_{12}(t) h_2^{-1}(t) \right] y_2(t) d\Lambda_{12}(t), \end{aligned} \quad (\text{A.4})$$

where

$$\begin{aligned} e_{jk}(t) &= I(j=1) - G_{1k}(t)/S_k(t), \\ v_{jk}(t) &= [d_{jk}(t) - e_{jk}(t)c_k(t)] h_k^{-1}(t). \end{aligned}$$

Finally a consistent estimator of $\sigma_{12}^{(1)}$ is obtained by replacing each unknown quantity in (9) by its consistent sample estimate. \square

Proof for Theorem 2. First, we derive the asymptotic joint distribution of $n^{-1/2}(\mathbf{U}_1(\boldsymbol{\beta}_1), \tilde{\mathbf{U}}_1(\boldsymbol{\gamma}_1))^T$. It can be shown that

$$\begin{aligned} n^{-1/2}\mathbf{U}_1(\boldsymbol{\beta}_1) &= n^{-1/2} \sum_{i=1}^n \mathbf{U}_{i1}(\boldsymbol{\beta}_1) + o_p(1), \\ n^{-1/2}\tilde{\mathbf{U}}_1(\boldsymbol{\gamma}_1) &= n^{-1/2} \sum_{i=1}^n (\boldsymbol{\eta}_i(\boldsymbol{\gamma}_1) + \boldsymbol{\phi}_i(\boldsymbol{\gamma}_1)) + o_p(1), \end{aligned} \quad (\text{A.5})$$

where

$$\begin{aligned} \mathbf{U}_{i1}(\boldsymbol{\beta}_1) &= \int_0^\infty \left\{ \mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right\} dM_{i1}(t), \\ \boldsymbol{\eta}_i &= \int_0^\infty \left\{ \mathbf{Z}_i^{(2)}(t) - \bar{\mathbf{z}}^{(2)}(\boldsymbol{\gamma}_1, t) \right\} w_i(t) d\tilde{M}_{i1}(t), \\ \boldsymbol{\phi}_i &= \int_0^\infty \frac{\mathbf{q}(t)}{\pi(t)} dM_i^c(t), \\ \tilde{M}_{i1}(t) &= \tilde{N}_{i1}(t) - \int_0^t \tilde{Y}_i(t) \exp(\boldsymbol{\gamma}_1^T \mathbf{Z}_i^{(2)}(u)) d\tilde{\Lambda}_{10}(u), \end{aligned}$$

$$M_i^c(t) = I(X_i \leq t, \delta_i = 0) - \int_0^t I(X_i \geq u) d\Lambda^c(u),$$

$$\mathbf{q}(t) = -n^{-1} \sum_{i=1}^n \int_0^\infty \left\{ \mathbf{Z}_i^{(2)}(u) - \bar{\mathbf{Z}}^{(2)}(\gamma_1, u) \right\} w_i(u) d\tilde{M}_{i1}(u) I(u \geq t > X_i),$$

and

$$\pi(t) = n^{-1} \sum_{i=1}^n I(X_i \geq t),$$

with $\tilde{\Lambda}_{10}(t) = \int_0^t \tilde{\lambda}_{10}(u) du$ being the baseline cause-specific cumulative hazard for cause 1,

$\Lambda^c(t) = \int_0^t \lambda^c(u) du$ the cumulative hazard for censoring variable,

$$\bar{\mathbf{z}}^{(1)}(\beta_1, t) = \frac{\lim_{n \rightarrow \infty} n^{-1} \sum_{l=1}^n Y_l(t) \mathbf{Z}_l^{(1)}(t) \exp(\beta_1^T \mathbf{Z}_l^{(1)}(t))}{\lim_{n \rightarrow \infty} n^{-1} \sum_{l=1}^n Y_l(t) \exp(\beta_1^T \mathbf{Z}_l^{(1)}(t))},$$

$$\text{and } \bar{\mathbf{z}}^{(2)}(\gamma_1, t) = \frac{\lim_{n \rightarrow \infty} \sum_{l=1}^n \omega_l(t) \tilde{Y}_l(t) \mathbf{Z}_l^{(2)}(t) \exp(\gamma_1^T \mathbf{Z}_l^{(2)}(t))}{\lim_{n \rightarrow \infty} \sum_{l=1}^n \omega_l(t) \tilde{Y}_l(t) \exp(\gamma_1^T \mathbf{Z}_l^{(2)}(t))}.$$

It follows from (A.5) and the multivariate central limit theorem that $n^{-1/2}(\mathbf{U}_1(\beta_1), \tilde{\mathbf{U}}_1(\gamma_1))^T$ has a $p + q$ multivariate normal limiting distribution with mean $\mathbf{0}$ and variance-covariance

$$\Omega^{(1)} = \begin{pmatrix} \Omega_{(pp)}^{(1)} & \Omega_{(pq)}^{(1)} \\ \Omega_{(qp)}^{(1)} & \Omega_{(qq)}^{(1)*} \end{pmatrix}, \quad (\text{A.6})$$

where

$$\Omega_{(pp)}^{(1)} = \int_0^\infty \left[\frac{\lim_{n \rightarrow \infty} n^{-1} \sum_l Y_l(t) \mathbf{Z}_l^{(1)}(t)^{\otimes 2} \exp(\beta_1^T \mathbf{Z}_l^{(1)}(t))}{\lim_{n \rightarrow \infty} n^{-1} \sum_l Y_l(t) \exp(\beta_1^T \mathbf{Z}_l^{(1)}(t))} - \bar{\mathbf{z}}^{(1)}(\beta_1, t)^{\otimes 2} \right] \lim n^{-1} \sum_{l=1}^n Y_l(t) \exp(\beta_1^T \mathbf{Z}_l^{(1)}(t)) d\Lambda_{10}(t), \quad (\text{A.7})$$

and

$$\Omega_{qq}^{*(1)} = E \left\{ (\boldsymbol{\eta}_i(\gamma_1) + \boldsymbol{\phi}_i(\gamma_1)) (\boldsymbol{\eta}_i(\gamma_1) + \boldsymbol{\phi}_i(\gamma_1))^T \right\}. \quad (\text{A.8})$$

Note that variance-covariance matrix between the two score test statistics is obtained as

the limit of

$$\begin{aligned}
& \left\langle n^{-1/2} \mathbf{U}_1(\boldsymbol{\beta}_1), n^{-1/2} \tilde{\mathbf{U}}_1(\boldsymbol{\gamma}_1) \right\rangle \\
&= n^{-1} \sum_{i=1}^n \langle \mathbf{U}_{i1}(\boldsymbol{\beta}_1), \boldsymbol{\eta}_i(\boldsymbol{\gamma}_1) + \boldsymbol{\phi}_i(\boldsymbol{\gamma}_1) \rangle \\
&= n^{-1} \sum_{i=1}^n \langle \mathbf{U}_{i1}(\boldsymbol{\beta}_1), \boldsymbol{\eta}_i(\boldsymbol{\gamma}_1) \rangle + n^{-1} \sum_{i=1}^n \langle \mathbf{U}_{i1}(\boldsymbol{\beta}_1), \boldsymbol{\phi}_i(\boldsymbol{\gamma}_1) \rangle \\
&= n^{-1} \sum_{i=1}^n \int_0^\infty \left\{ \mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right\} \left\{ \mathbf{Z}_i^{(2)}(t) - \bar{\mathbf{z}}^{(2)}(\boldsymbol{\gamma}_1, t) \right\} \omega_i(t) dt < M_{i1}, \tilde{M}_{i1} > (t) \\
&\quad + n^{-1} \sum_{i=1}^n \int_0^\infty \left\{ \mathbf{Z}_i^{(1)} - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right\} \frac{\mathbf{q}(t)}{\pi(t)} dt < M_{i1}, M_i^c > (t),
\end{aligned}$$

which converges in probability to

$$\begin{aligned}
\boldsymbol{\Omega}_{(pq)}^{(1)} &= E \int_0^\infty \left\{ \mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right\} \left\{ \mathbf{Z}_i^{(2)}(t) - \bar{\mathbf{z}}^{(2)}(\boldsymbol{\gamma}_1, t) \right\} \omega_i(t) dt < M_{i1}, \tilde{M}_{i1} > (t) \\
&\quad + E \int_0^\infty \left\{ \mathbf{Z}_i^{(1)} - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right\} \frac{\mathbf{q}(t)}{\pi(t)} dt < M_{i1}, M_i^c > (t).
\end{aligned} \tag{A.9}$$

Let $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\gamma}}_1$ be solutions to $\mathbf{U}_1(\hat{\boldsymbol{\beta}}_1) = 0$ and $\tilde{\mathbf{U}}_1(\hat{\boldsymbol{\gamma}}_1) = 0$, respectively. Applying Taylor series expansion to $(\mathbf{U}_1(\hat{\boldsymbol{\beta}}_1), \tilde{\mathbf{U}}_1(\hat{\boldsymbol{\gamma}}_1))^T$ around $(\boldsymbol{\beta}_1, \boldsymbol{\gamma}_1)$, we have

$$n^{1/2} \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 \\ \hat{\boldsymbol{\gamma}}_1 - \boldsymbol{\gamma}_1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Omega}_{(pp)}^{(1)-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{(qq)}^{(1)-1} \end{pmatrix} \begin{pmatrix} \mathbf{U}_1(\boldsymbol{\beta}_1) \\ \tilde{\mathbf{U}}_1(\boldsymbol{\gamma}_1) \end{pmatrix} + o_p(1),$$

where

$$\begin{aligned}
\boldsymbol{\Omega}_{(qq)}^{(1)} &= \int_0^\infty \left\{ \frac{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n \omega_l(t) \tilde{Y}_l(t) \mathbf{Z}_l^{(2)}(t)^{\otimes 2} \exp(\boldsymbol{\gamma}_1^T \mathbf{Z}_l^{(2)}(t))}{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n \omega_l(t) \tilde{Y}_l(t) \exp(\boldsymbol{\gamma}_1^T \mathbf{Z}_l^{(2)}(t))} - \bar{\mathbf{z}}^{(2)}(\boldsymbol{\gamma}_1, t)^{\otimes 2} \right\} \\
&\quad \lim_{n \rightarrow \infty} n^{-1} \sum_{l=1}^n \omega_l(t) \tilde{Y}_l(t) \exp(\boldsymbol{\gamma}_1^T \mathbf{Z}_l^{(2)}(t)) d\tilde{\Lambda}_{10}(t).
\end{aligned} \tag{A.10}$$

This, together with (A.6), implies that

$$\boldsymbol{\Sigma}^{(1)} = \begin{pmatrix} \boldsymbol{\Omega}_{(pp)}^{(1)-1} & \boldsymbol{\Omega}_{(pp)}^{(1)-1} \boldsymbol{\Omega}_{(pq)}^{(1)} \boldsymbol{\Omega}_{(qq)}^{(1)-1} \\ \boldsymbol{\Omega}_{(qq)}^{(1)-1} \boldsymbol{\Omega}_{(qp)}^{(1)} \boldsymbol{\Omega}_{(pp)}^{(1)-1} & \boldsymbol{\Omega}_{(qq)}^{(1)-1} \boldsymbol{\Omega}_{(qq)}^{(1)*} \boldsymbol{\Omega}_{(qq)}^{(1)-1} \end{pmatrix}. \tag{A.11}$$

A consistent estimator for $\boldsymbol{\Sigma}^{(1)}$ is obtained by replacing all unknown quantities with their respective sample estimates in (21) in section 3. \square

Proof for Theorem 3. Under the null hypothesis, it was shown by Fleming and Harrington (1991) that

$$\begin{aligned}\mathbf{U}_1(\boldsymbol{\beta}_1) &= \sum_{i=1}^n \int_0^\tau \left(\mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) dM_{i1}(t) - n \int_0^\tau \left(\bar{\mathbf{Z}}^{(1)}(\boldsymbol{\beta}_1, t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) dM_{i1}(t), \\ \mathbf{U}(\boldsymbol{\beta}) &= \sum_{i=1}^n \int_0^\tau \left(\mathbf{Z}_i^{(3)}(t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}, t) \right) dM_i(t) - n \int_0^\tau \left(\bar{\mathbf{Z}}^{(3)}(\boldsymbol{\beta}, t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}, t) \right) dM_i(t)\end{aligned}$$

where

$$\begin{aligned}M_{i1}(t) &= N_{i1}(t) - \int_0^t \lambda_{j0}(u) Y_i(u) \exp(\boldsymbol{\beta}_1^T \mathbf{Z}_i^{(1)}(u)) du, \\ M_i(t) &= N_i(t) - \int_0^t \lambda_0(u) Y_i(u) \exp(\boldsymbol{\beta}^T \mathbf{Z}_i^{(3)}(u)) du,\end{aligned}$$

$$\text{and } \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}, t) = \frac{\lim_{n \rightarrow \infty} n^{-1} \sum_{l=1}^n Y_l(t) \mathbf{Z}_l^{(3)}(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_l^{(3)}(t))}{\lim_{n \rightarrow \infty} n^{-1} \sum_{l=1}^n Y_l(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_l^{(3)}(t))}.$$

The first part of $(\mathbf{U}_1(\boldsymbol{\beta}_1), \mathbf{U}(\boldsymbol{\beta}))^T, i = 1, 2, \dots, n$ can be viewed as a sum of independently identically distributed random vector. By using multivariate central limit theory, we can prove the first part of the vector (2) has a bivariate normal distribution with mean $\mathbf{0}$, and variance-covariance matrix $\boldsymbol{\Omega}^{(2)}$. Since $\bar{\mathbf{Z}}^{(1)}(\boldsymbol{\beta}_1, t)$ and $\bar{\mathbf{Z}}^{(3)}(\boldsymbol{\beta}, t)$ converge in probability to some deterministic process $\bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t)$ and $\bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}, t)$, respectively, we can prove the second part of the vector (2) converge in probability to zero by using the central limit theory for stochastic integrals with respect to counting process martingales. Then we can use Slutsky theorem to prove $n^{-1/2}(\mathbf{U}_1(\boldsymbol{\beta}_1), \mathbf{U}(\boldsymbol{\beta}))^T$ has a $p + q$ dimension multivariate normal limiting distribution with mean $\mathbf{0}$ and variance-covariance matrix

$$\begin{pmatrix} \boldsymbol{\Omega}_{(pp)}^{(2)} & \boldsymbol{\Omega}_{(pq)}^{(2)} \\ \boldsymbol{\Omega}_{(qp)}^{(2)} & \boldsymbol{\Omega}_{(qq)}^{(2)} \end{pmatrix}, \quad (\text{A.12})$$

where $\boldsymbol{\Omega}_{(pp)}^{(2)}$ is defined in (A.7),

$$\begin{aligned}\boldsymbol{\Omega}_{(qq)}^{(2)} &= \int_0^\infty \left[\frac{\lim_{n \rightarrow \infty} n^{-1} \sum_{l=1}^n Y_l(t) \mathbf{Z}_l^{(3)}(t)^{\otimes 2} \exp(\boldsymbol{\beta}^T \mathbf{Z}_l^{(3)}(t))}{\lim_{n \rightarrow \infty} n^{-1} \sum_{l=1}^n Y_l(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_l^{(3)}(t))} - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}, t)^{\otimes 2} \right] \\ &\quad \lim n^{-1} \sum_{l=1}^n Y_l(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_l^{(3)}(t)) d\Lambda_0(t).\end{aligned}$$

and the covariance

$$\Omega_{(pq)}^{(2)} = E \int_0^\tau \left(\mathbf{Z}_i^{(3)}(t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}_., t) \right) \left(\mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) \exp(\boldsymbol{\beta}_1^T \mathbf{Z}_i^{(1)}(t)) Y_i(t) d\Lambda_{10}(t)$$

is the limit of

$$\begin{aligned} & \langle n^{-1/2} \mathbf{U}_.(\boldsymbol{\beta}_.), n^{-1/2} \mathbf{U}_1(\boldsymbol{\beta}_1) \rangle \\ &= n^{-1} \sum_{i=1}^n \left\langle \int_0^\tau \left(\mathbf{Z}_i^{(3)}(t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}_., t) \right) dM_i(t), \int_0^\tau \left(\mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) dM_{i1}(t) \right\rangle \\ &= n^{-1} \sum_{i=1}^n \int_0^\tau \left(\mathbf{Z}_i^{(3)}(t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}_., t) \right) \left(\mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) d < M_i, M_{i1} > (t) \\ &= n^{-1} \sum_{i=1}^n \int_0^\tau \left(\mathbf{Z}_i^{(3)}(t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}_., t) \right) \left(\mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) d < M_{i1+i2}, M_{i1} > (t) \\ &= n^{-1} \sum_{i=1}^n \int_0^\tau \left(\mathbf{Z}_i^{(3)}(t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}_., t) \right) \left(\mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) d < M_{i1}, M_{i1} > (t) \\ &= n^{-1} \sum_{i=1}^n \int_0^\tau \left(\mathbf{Z}_i^{(3)}(t) - \bar{\mathbf{z}}^{(3)}(\boldsymbol{\beta}_., t) \right) \left(\mathbf{Z}_i^{(1)}(t) - \bar{\mathbf{z}}^{(1)}(\boldsymbol{\beta}_1, t) \right) \exp(\boldsymbol{\beta}_1^T \mathbf{Z}_i^{(1)}(t)) Y_i(t) d\Lambda_{10}(t). \end{aligned}$$

Applying the Taylor series expansion to $(\mathbf{U}_1(\hat{\boldsymbol{\beta}}_1), \mathbf{U}_.(\hat{\boldsymbol{\beta}}_.))^T$ around $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_.)$, we have

$$n^{1/2} \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 \\ \hat{\boldsymbol{\beta}}_. - \boldsymbol{\beta}_. \end{pmatrix} \approx \begin{pmatrix} \boldsymbol{\Omega}_{(pp)}^{(2)-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{(qq)}^{(2)-1} \end{pmatrix} \begin{pmatrix} \mathbf{U}_1(\boldsymbol{\beta}_1) \\ \mathbf{U}_.(\boldsymbol{\beta}_.) \end{pmatrix}.$$

This, together with (A.12), implies that

$$\boldsymbol{\Sigma}^{(2)} = \begin{pmatrix} \boldsymbol{\Omega}_{(pp)}^{(2)-1} & \boldsymbol{\Omega}_{(pp)}^{(2)-1} \boldsymbol{\Omega}_{(pq)}^{(2)} \boldsymbol{\Omega}_{(qq)}^{(2)-1} \\ \boldsymbol{\Omega}_{(qq)}^{(2)-1} \boldsymbol{\Omega}_{qp}^{(2)} \boldsymbol{\Omega}_{(pp)}^{(2)-1} & \boldsymbol{\Omega}_{(qq)}^{(2)-1} \end{pmatrix}. \quad (\text{A.13})$$

Finally, a consistent estimator for $\boldsymbol{\Sigma}^{(2)}$ is obtained by replacing all unknown quantities with their respective sample estimates in (A.13). \square

Derivation for Covariance between U_{1k} and $\tilde{U}_{1k'}$

$$\begin{aligned}
& n^{-1/2} < (U_{1k}, \tilde{U}_{1k'}) > \\
& = < \int_0^\tau W_1(t) Y_K(t) \left(\frac{dN_{1k}(t)}{Y_k(t)} - \frac{dN_{1\cdot}(t)}{Y_{\cdot}(t)} \right), \sum_{l=1}^K \sum_{j=1}^2 (A_{jk'l}(\tau) + c_{k'l}(\tau) B_{jl}(\tau)) > \\
& = < \int_0^\tau W_1(t) Y_k(t) \left(\frac{dM_{1k}(t)}{Y_k(t)} - \frac{dM_{1\cdot}(t)}{Y_{\cdot}(t)} \right), \\
& \quad \sum_{l=1}^K \int_0^\tau V_{1k'l}(t) dM_{1l}(t) + c_{k'l}(\tau) \int_0^\tau E_{1l}(t) \hat{h}_l^{-1}(t) dM_{1l}(t) > \\
& = < \int_0^\tau W_1(t) Y_k(t) \frac{dM_{1k}(t)}{Y_k(t)}, \int_0^\tau V_{1k'k}(t) dM_{1k}(t) + c_{k'k}(\tau) \int_0^\tau E_{1k}(t) \hat{h}_k^{-1}(t) dM_{1k}(t) > \\
& + \int_0^\tau W_1(t) Y_k(t) \frac{\sum_{l=1}^K dM_{1l}(t)}{Y_{\cdot}(t)}, \sum_{l=1}^K \int_0^\tau V_{1k'l}(t) dM_{1l}(t) + c_{k'l}(\tau) \int_0^\tau E_{1l}(t) \hat{h}_l^{-1}(t) dM_{1l}(t) > \\
& = \left(\int_0^\tau W_1(t) V_{1k'k}(t) + c_{k'k}(\tau) \int_0^\tau W_1(t) E_{1k}(t) \hat{h}_k^{-1}(t) \right) d < M_{1k}(t), M_{1k}(t) > \\
& + \sum_{l=1}^K \left(\int_0^\tau W_1(t) \frac{Y_k(t)}{Y_{\cdot}(t)} V_{1k'l}(t) + c_{k'l}(\tau) \int_0^\tau W_1(t) \frac{Y_k(t)}{Y_{\cdot}(t)} E_{1l}(t) \hat{h}_l^{-1}(t) \right) d < M_{1l}(t), M_{1l}(t) > \\
& = \left(\int_0^\tau W_1(t) V_{1k'k}(t) + c_{k'k}(\tau) \int_0^\tau W_1(t) E_{1k}(t) \hat{h}_k^{-1}(t) \right) Y_k(t) d\Lambda_{1k}(t) \\
& + \sum_{l=1}^K \left(\int_0^\tau W_1(t) \frac{Y_k(t)}{Y_{\cdot}(t)} V_{1k'l}(t) + c_{k'l}(\tau) \int_0^\tau W_1(t) \frac{Y_k(t)}{Y_{\cdot}(t)} E_{1l}(t) \hat{h}_l^{-1}(t) \right) Y_l(t) d\Lambda_{1l}(t),
\end{aligned}$$

where $V_{jkl}(t) = [D_{jkl}(t) - E_{jl}(t)c_{kl}(t)] \hat{h}_l^{-1}(t)$ and all other quantities are defined in Gray (1988) on page 1153. $n^{-1/2} < (U_{1k}, \tilde{U}_{1k'}) >$ converges in probability to

$$\begin{aligned}
& cov(n^{-1/2} U_{1k}, n^{-1/2} \tilde{U}_{1k'}) \\
& = \left(\int_0^\tau w_1(t) v_{1k'k}(t) + c_{k'k}(\tau) \int_0^\tau w_1(t) e_{1k}(t) \hat{h}_k^{-1}(t) \right) y_k(t) d\Lambda_{1k}(t) \\
& + \sum_{l=1}^K \left(\int_0^\tau w_1(t) \frac{y_k(t)}{y_{\cdot}(t)} v_{1k'l}(t) + c_{k'l}(\tau) \int_0^\tau w_1(t) \frac{y_k(t)}{y_{\cdot}(t)} e_{1l}(t) \hat{h}_l^{-1}(t) \right) y_l(t) d\Lambda_{1l}(t),
\end{aligned} \tag{A.14}$$

where

$$\begin{aligned}
e_{jk}(t) &= I(j=1) - G_{1k}(t)/S_k(t) \\
v_{jkl}(t) &= [d_{jkl}(t) - e_{jl}(t)c_{kl}(t)] \hat{h}_l^{-1}(t),
\end{aligned}$$

and d_{jkl} is defined in Gray (1988) on page 1146. Finally a consistent estimator of $cov(n^{-1/2} U_{1k}, n^{-1/2} \tilde{U}_{1k'})$ is obtained by replacing each unknown quantity in (A.14) by its consistent sample estimate.

□

APPENDIX A.2. Additional Simulation Results

Simulation results for one-sided two-sample tests with respect to the CSH and CIF pair under the simulation setting of Figure 1

Under the simulation setting of Figure 1, we also conducted a simulation for the one-sided two-sample tests with respect to the CSH and CIF pair. The results are presented in Figure A.1 below.

Simulation results for two-sample comparisons with respect to the CSH and ACH pair

Below we present a simulation for the two-group comparison problem with respect to the CSH and all-cause hazard (ACH) pair. Let λ_{1k} and λ_k denote the CSH for type 1 failure and the ACH, respectively, for group k ($k = 1, 2$). Assume that in each group, both types of failures have constant cause-specific hazards and thus the all-cause hazard is also constant. The censoring rate is set to be 0.1 with an independent exponential censoring time. The nominal significance level is 0.05.

Figure A.2 below depicts the simulated rejection power curves of the two-sided chi-square joint test, maximum joint test, and Bonferroni joint test for (11) under four scenarios. Figure A.2(a) represents a null case where there is no difference with respect to type 1 failure between the two groups ($\lambda_{11} = \lambda_{12} = 0.04, \lambda_{1\cdot} = \lambda_{2\cdot} = 0.05$). Figure A.2(b) corresponds to a situation where the group difference in CSH is smaller than ACH ($\lambda_{11} = 0.6, \lambda_{12} = 0.61, \lambda_{2\cdot} = 0.7, \lambda_{2\cdot} = 0.81$). Figure A.2(c) corresponds to a situation where the group difference in CSH is bigger than ACH ($\lambda_{11} = 0.05, \lambda_{12} = 0.0625, \lambda_{2\cdot} = 0.058, \lambda_{2\cdot} = 0.17$).

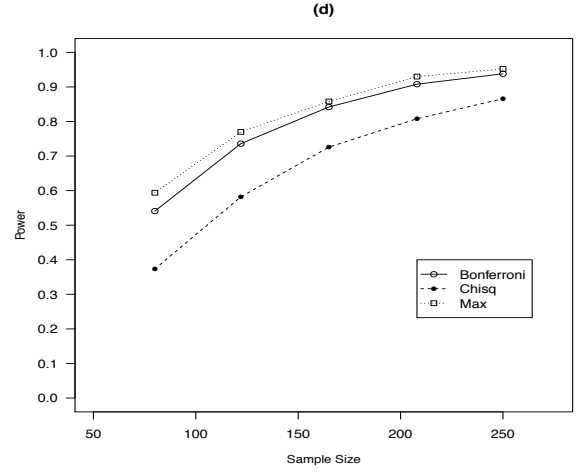
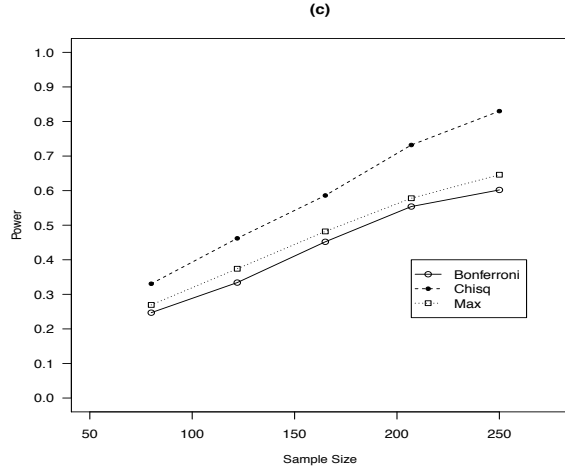
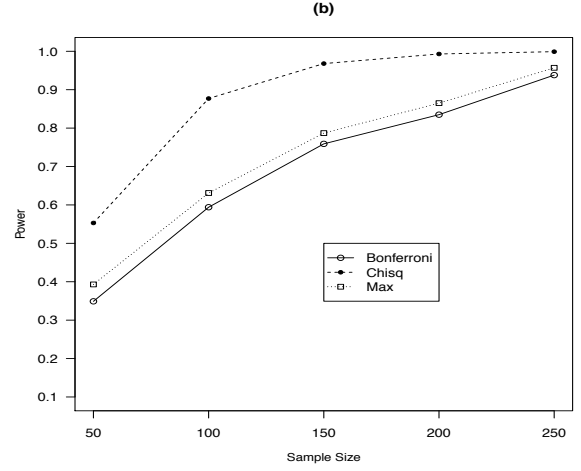
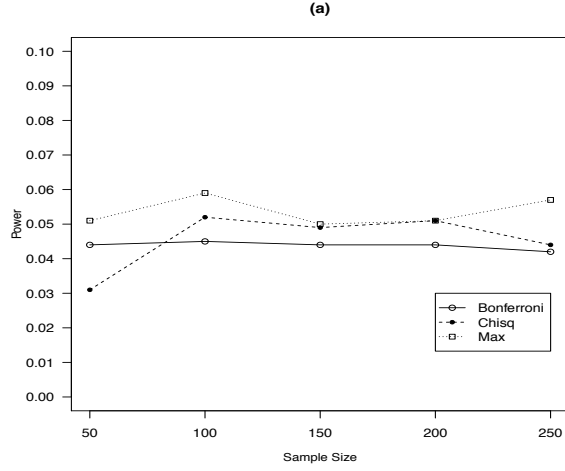


Figure A.1: Simulated power of the one-sided chi-square joint test, maximum joint test and Bonferroni joint test for two-group type 1 failure comparison with respect to the CSH and CIF pair under four scenarios as described in Section 4: (a) null case under H_0 , (b) smaller group difference in CSH and larger group difference in CIF, (b) larger group difference in CSH and smaller group difference in CIF, and (d) similar group effects on CSH and CIF.

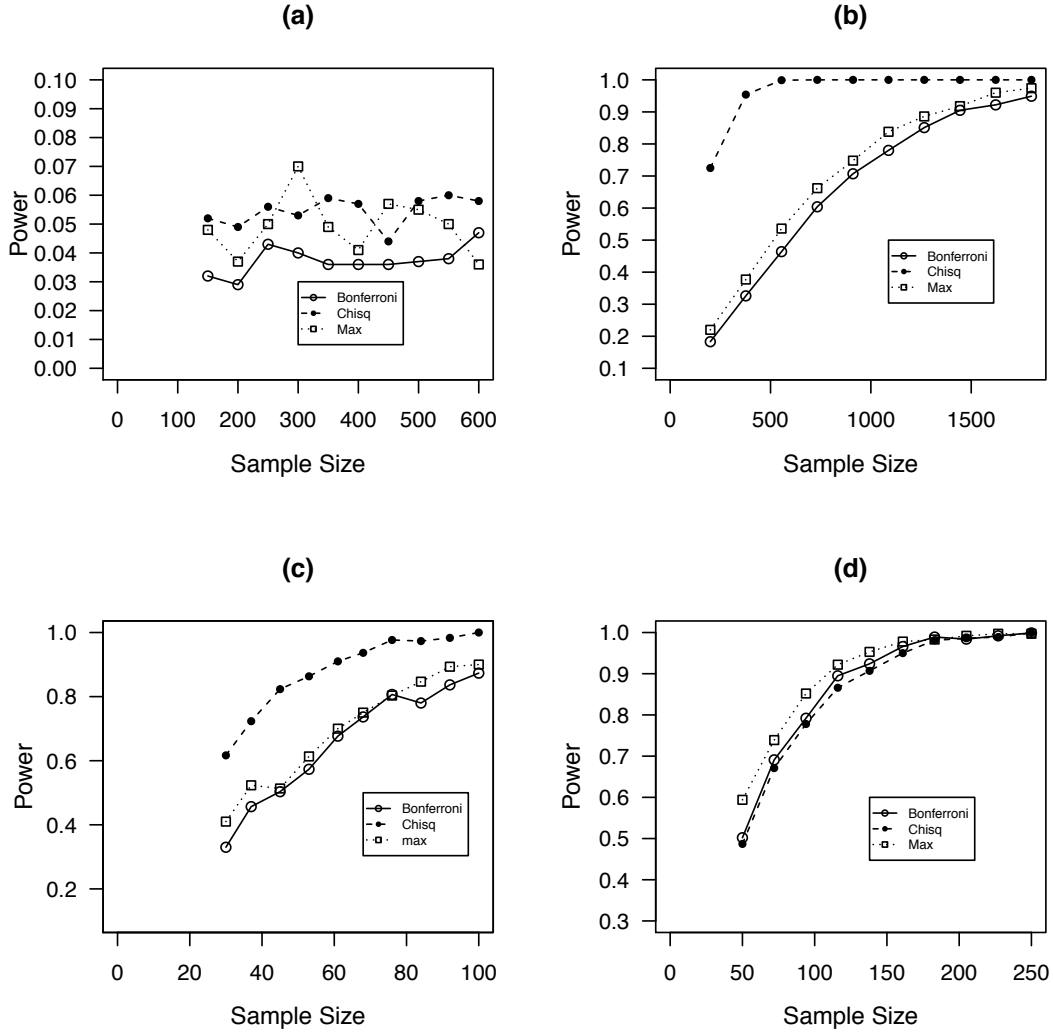


Figure A.2: Simulated power of the two-sided chi-square joint test, maximum joint test and Bonferroni joint test for two-group type 1 failure comparison with respect to the CSH and ACH pair under four scenarios: (a) null case under H_0 , (b) smaller group difference in CSH and larger group difference in ACH, (b) larger group difference in CSH and smaller group difference in ACH, and (d) similar group effects on CSH and ACH.

Figure A.3 shows the power curves for the three tests under the same scenarios as Figure A.2, except that one-sided tests are performed.

It is seen from Figure A.2 and Figure A.3 that the results for the (CSH, ACH) pair is consistent with those for the (CSH, CIF) pair. The type I errors are well controlled at nominal level 0.05 for both two-sided and one-sided tests. When the group differences are different in the two quantities, the chi-square joint test is much more powerful than the Bonferroni test and the maximum test. When the group differences are similar in CSH and ACH, then the maximum joint test performs better especially for a one-sided test.

Simulation results for joint regression of CSH and ACH

This section presents some simulation results for the joint Cox model with respect to the CSH and ACH pair described in Section 3.2. Assume the following model:

$$\begin{aligned}\lambda_1(t|\mathbf{Z}) &= \lambda_{10}(t) \exp(\beta_{11}Z_1 + \beta_{21}Z_2) \\ \lambda(t|\mathbf{Z}) &= \lambda_0(t) \exp(\beta_1.Z_1 + \beta_2.Z_2),\end{aligned}\tag{A.15}$$

where Z_1 and Z_2 are binary variables. We are interested in testing the effects of Z_1 on both CSH and ACH, or $H_0 : \beta_{11} = 0$ and $\beta_{1.} = 0$. We generated data under various different alternatives. Figure A.4 (a) represents the null case where $\beta_{11} = \beta_{1.} = 0$. Figure A.4 (b) corresponds to a situation where the effect size of Z_1 for CSH is smaller than ACH ($H_a : \beta_{11} = 0$ and $\beta_{1.} = -0.13$). Figure A.4(c) corresponds to a situation where the effect size of Z_1 for CSH is bigger than ACH ($H_a : \beta_{11} = -0.15$ and $\beta_{1.} = 0$). Figure A.4(d) corresponds to a case when the effect sizes of Z_1 are similar for CSH and ACH ($H_a : \beta_{11} = -0.15$ and $\beta_{1.} = -0.14$). For all the four scenarios, $\lambda_{10}(t) = 0.06$, $\lambda_0(t) = 0.08$, $\beta_{21} = 0.3$ and $\beta_2 = 0.24$.

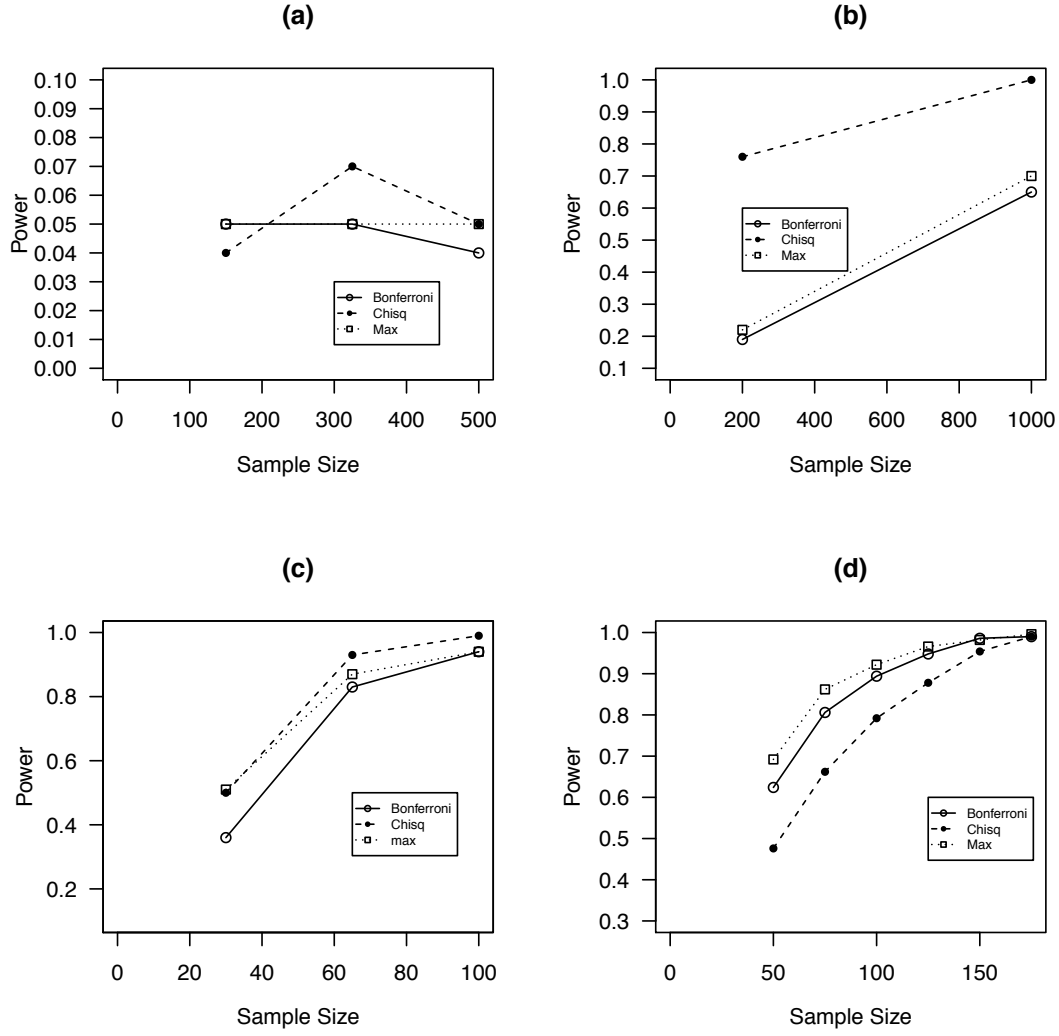


Figure A.3: Simulated power of the one-sided chi-square joint test, maximum joint test and Bonferroni joint test for two-group type 1 failure comparison with respect to the CSH and ACH pair under four scenarios: (a) null case under H_0 , (b) smaller group difference in CSH and larger group difference in ACH, (b) larger group difference in CSH and smaller group difference in ACH, and (d) similar group effects on CSH and ACH.

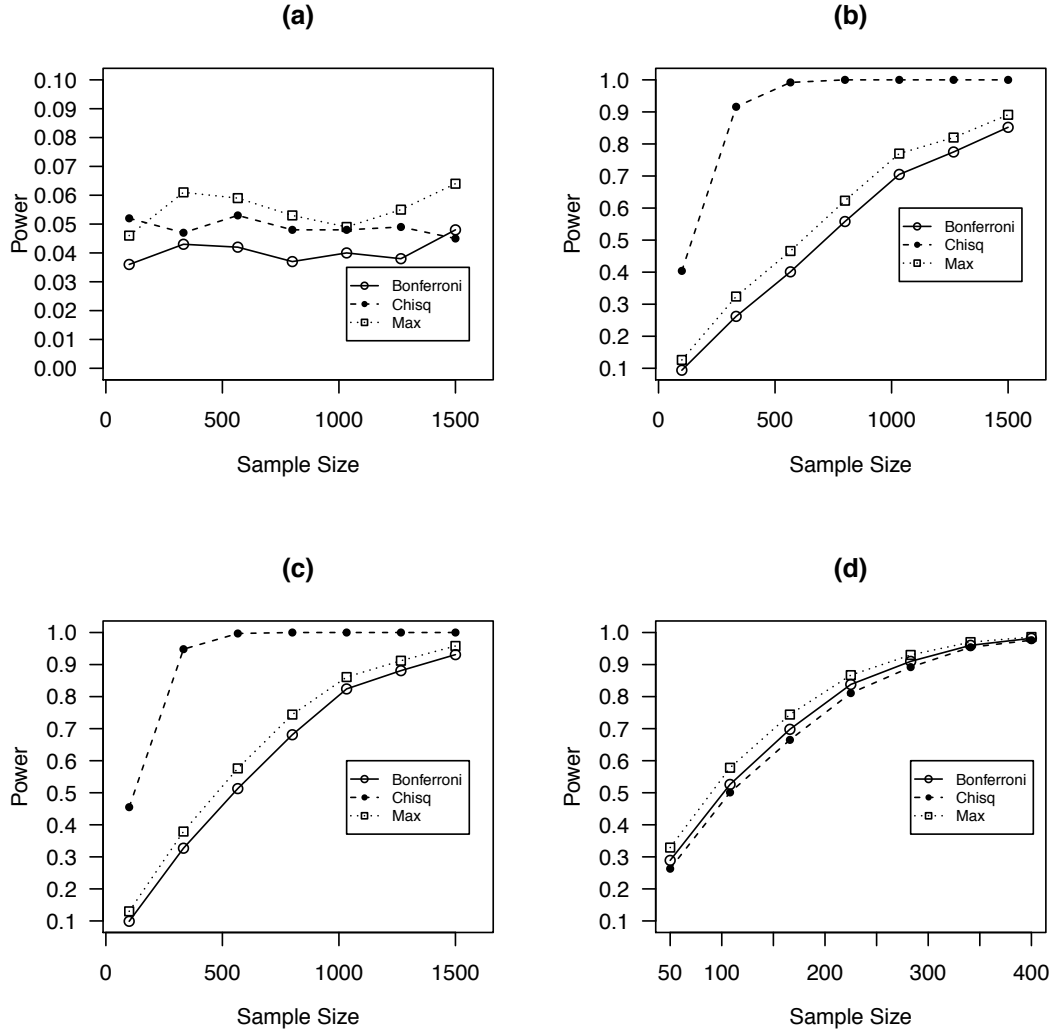


Figure A.4: Simulated power of the two-sided chi-square joint test, maximum joint test and Bonferroni joint test of a local hypothesis ($H_0 : \beta_{11} = 0$ and $\beta_{1\cdot} = 0$.) for a joint regression model (A.15) of CSH and ACH under four scenarios: (a) null case, (b) smaller effects on CSH and larger effects on ACH, (b) larger effects on CSH and smaller effects on ACH, and (d) similar effects on CSH and ACH.

Figure A.4 shows that the results are consistent with what we have observed for the CSH and CIF pair.

APPENDIX A.3. Additional Simulation Results

Graphical display of the cumulative incidence function $F_1(t)$ by group under the simulation setting of Figure 1

The cumulative incidence function $F_1(t)$ by group under the simulation setting of Figure 1 is illustrated in Figure A.5 below.

Cox-Snell residual plots for Follicular cell lymphoma study

For the Follicular cell lymphoma data, we constructed the Cox-Snell plot to check the overall fit of the Cox model for the all-cause hazard and the cause-specific hazard. The plots, along with pointwise 95% bootstrap confidence intervals are depicted in Figure A.6 below.

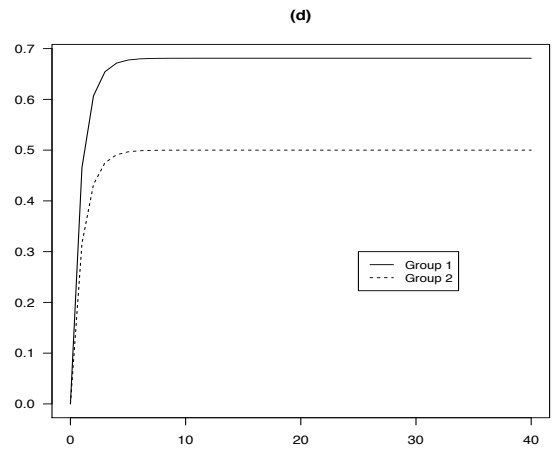
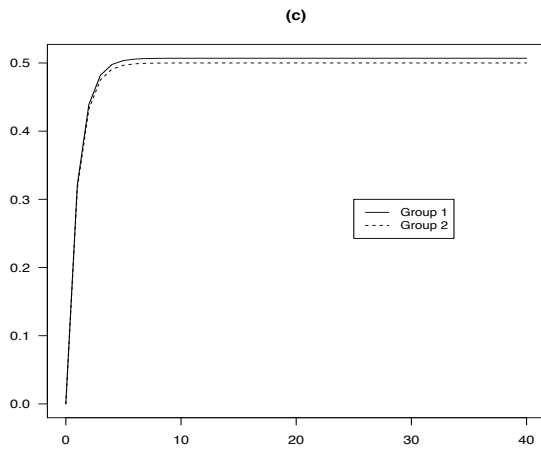
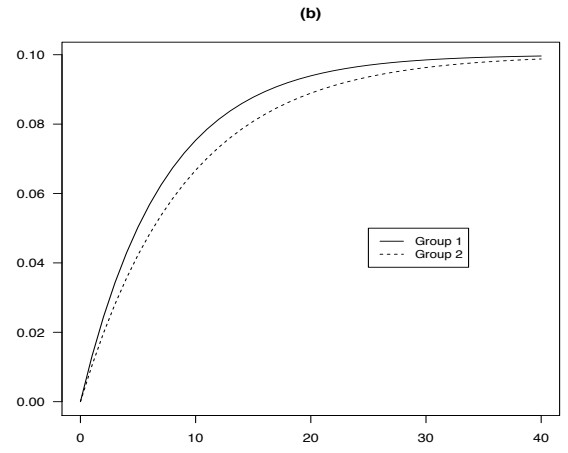
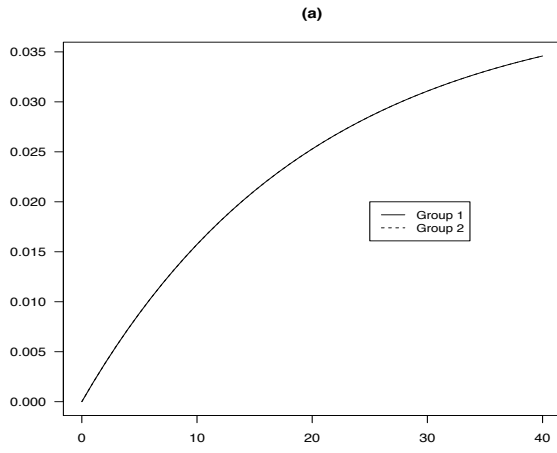


Figure A.5: Graphical illustration of the cumulative incidence function (CIF) by group under the simulation setting of Figure 1.

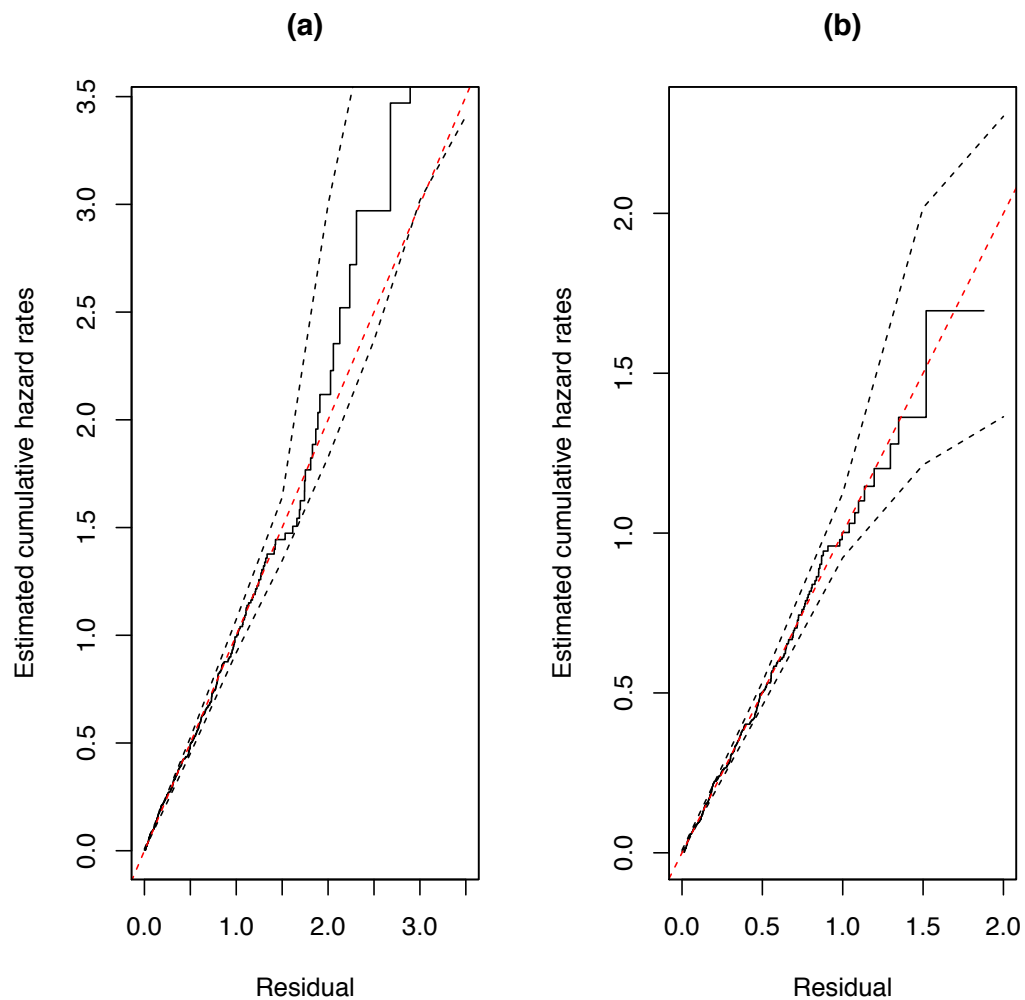


Figure A.6: The Cox-Snell residual plot (solid line) for the proportional all-cause hazards model (panel (a)) and the proportional cause-specific hazards model (panel(b)), with point-wise 95% bootstrap confidence intervals (dashed lines), and the 45 degree line (dotted lines) for the Follicular cell lymphoma data