Supplemental material to

A LOGRANK TEST-BASED METHOD FOR SIZING CLINICAL TRIALS WITH TWO CO-PRIMARY TIME-TO-EVENTS ENDPOINTS

TOMOYUKI SUGIMOTO*
Department of Mathematical Sciences, Hirosaki University,
3 Bunkyo-cho, Hirosaki, Aomori, 036-8097, Japan
e-mail: tomoyuki@cc.hirosaki-u.ac.jp

TAKASHI SOZU
Department of Biostatistics,
Kyoto University School of Public Heath, Japan

TOSHIMITSU HAMASAKI
Department of Biomedical Statistics,
Osaka University Graduate School of Medicine, Japan

SCOTT EVANS
Department of Biostatistics,
Harvard University School of Public Heath, USA

*Corresponding author

Notations used in these supplementary materials are taken from the main text, in particular, as long as their explanations are not newly provided here.

A Joint survival modelling adopted in the article

A.1 Three copula models: Brief characteristics

In the present article, we consider the three typical models belonging to the Archimedean copula model family in order to model the joint survival function by

$$S^{(k)}(t,s;\theta) = \mathcal{C}(S_1^{(k)}(t), S_2^{(k)}(s);\theta), \ \ k=1,2.$$

We here supplement the explanation on these models. For simplicity, suppose $g_i = k$ (the i^{th} participant belongs to the group k) and hence $\rho^{(k)}$ is the correlation for the i^{th} bivariate survival times.

Clayton copula. The bivariate survival function given by Clayton (1978) is

$$\mathcal{C}(u,v;\theta) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}, \ \theta \geq 0.$$

Although this model allows negative dependencies if $\theta < 0$, we do not consider the case. If $\theta = \theta^{(k)} > 0$, then the bivariate survival times T_{i1}^* and T_{i2}^* have a positive correlation $\rho^{(k)}$. The case of $\theta^{(k)} = 0$ provides the independence between T_{i1}^* and T_{i2}^*. As $\theta^{(k)} \to \infty$, this achieves the upper Fréchet bound $S^{(k)}(t,s) = \min(S_1^{(k)}(t),\ S_2^{(k)}(s))$, $k=1,2$.

Gumbel copula. The bivariate function induced from a positive stable distribution is

$$\mathcal{C}(u,v;\theta) = \exp\left(-\left\{(-\log u)^{1/\theta} + (-\log v)^{1/\theta}\right\}^{\theta}\right), \ 0 \leq \theta \leq 1.$$

A smaller value of $\theta = \theta^{(k)}$ gives a larger positive correlation between T_{i1}^* and T_{i2}^*. When $\theta^{(k)} = 1$, T_{i1}^* and T_{i2}^* are mutually independent (Hougaard, 1986; Shih and Louis, 1995).

Frank copula. The bivariate function introduced by Frank (1979) is

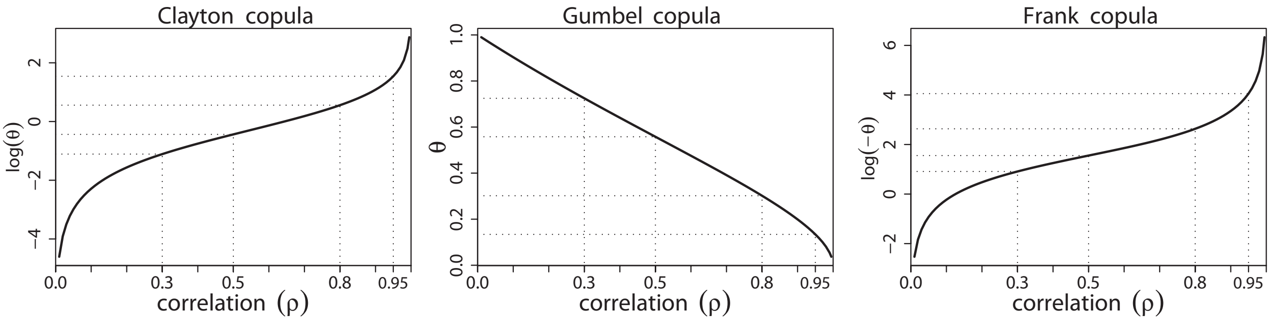
$$\mathcal{C}(u,v;\theta) = \theta^{-1}\log\left(1 + \frac{(\mathrm{e}^{\theta u}-1)(\mathrm{e}^{\theta v}-1)}{\mathrm{e}^{\theta}-1}\right).$$

When $\theta = \theta^{(k)} = 0$, T_{i1}^* and T_{i2}^* are mutually independent. If $\theta^{(k)} < 0$, then T_{i1}^* and T_{i2}^* are positively correlated. Note that, originally and in many articles, this copula model is usually parameterized on e^{θ}. We adopt the parametrization as displayed, because it is easier to treat our computational problems on θ than the exponential scale. One computational problem in this copula model is that the inside $(\mathrm{e}^{\theta u}-1)(\mathrm{e}^{\theta v}-1)/(\mathrm{e}^{\theta}-1)$ of the logarithm may have a serious rounding error when e^{θ} is near zero, such as $\theta < -\mathrm{e}^2$. Then, we could counter this problem with an approximation

$$\frac{(\mathrm{e}^{\theta u}-1)(\mathrm{e}^{\theta v}-1)}{(\mathrm{e}^{\theta}-1)} \approx -(1+\mathrm{e}^{\theta}+\mathrm{e}^{2\theta})(1-\mathrm{e}^{\theta u}-\mathrm{e}^{\theta v}+\mathrm{e}^{\theta(u+v)}).$$

These copula models have different characteristics of the bivariate dependence. Generally, the Clayton and Gumbel copulas provide late and early dependences, respectively, and the Frank copula describes a symmetric dependence without tail dependence.

Figure 1 gives the relationships between the correlation $\rho^{(k)}$ and a transformation of $\theta^{(k)}$ for the three copula models in bivariate continuous survival data, from which a list of $\theta^{(k)}$ to $\rho^{(k)} = 0.3, 0.5, 0.8, 0.95$ is picked.



Copula	Association form	$\rho^{(k)} = 0.3$	$\rho^{(k)} = 0.5$	$\rho^{(k)} = 0.8$	$\rho^{(k)} = 0.95$
Clayton	$\log(\theta^{(k)})$	-1.1158	-0.4439	0.5512	1.5313
Gumbel	$\theta^{(k)}$	0.7249	0.5582	0.3027	0.1357
Frank	$\log(-\theta^{(k)})$	0.9116	1.5539	2.6350	4.0579

Figure 1: Relationships between $\theta^{(k)}$ and $\rho^{(k)}$ for the three copula models. The y-axes of plots for the Clayton and Frank copulas are drawn in $\log(\theta^{(k)})$ and $\log(-\theta^{(k)})$.

A.2 A method to compute the association parameter from correlation

For simplicity, the superscripts (k) for $\theta^{(k)}$ and $\rho^{(k)}$ are omited. We here introduce a method to calculate the corresponding θ to a specified ρ. Figure 1 is created using this algorithm.

The Newton-Raphson (NR) algorithm is performed to numerically find θ which satisfies the equation $\tilde{\rho}(\theta) - \rho = 0$, where

$$\tilde{\rho}(\theta) = \int_0^\infty \int_0^\infty \mathcal{C}\left(\mathrm{e}^{-t}, \mathrm{e}^{-s}; \theta\right) dt ds - 1 \quad (A.1)$$

$$= \int_0^1 \int_0^1 \mathcal{C}\left(u, v; \theta\right) u^{-1} v^{-1} du dv - 1 \quad (A.2)$$

The latter equation (A.2) may be preferred because of not including infinite intervals. The first derivative needed for the NR method is

$$\frac{d}{d\theta}\tilde{\rho}(\theta) = \int_0^\infty \int_0^\infty \frac{d}{d\theta}\mathcal{C}\left(\mathrm{e}^{-t}, \mathrm{e}^{-s}; \theta\right) dt ds \quad (A.3)$$

$$= \int_0^1 \int_0^1 \frac{d}{d\theta}\mathcal{C}\left(u, v; \theta\right) u^{-1} v^{-1} du dv, \quad (A.4)$$

where

$$\frac{d\mathcal{C}\left(u, v; \theta\right)}{d\theta} = \begin{cases} \mathcal{C}\left(u, v; \theta\right) \theta^{-1} \left\{ \frac{v^\theta \log(u) + u^\theta \log(v)}{u^\theta + v^\theta - u^\theta v^\theta} - \log(\mathcal{C}\left(u, v; \theta\right)) \right\} & \text{(Clayton copula)} \\ \mathcal{C}\left(u, v; \theta\right) (\phi_u + \phi_v)^{\theta - 1} \{ \phi_u \log(\phi_u) + \phi_v \log(\phi_v) & \\ \qquad - (\phi_u + \phi_v) \log(\phi_u + \phi_v) \} & \text{(Gumbel copula)} \\ \frac{u\phi_u(\phi_v - 1) + v\phi_v(\phi_u - 1) - (\phi_u - 1)(\phi_v - 1)\mathrm{e}^\theta / (\mathrm{e}^\theta - 1)}{\theta \{ (\phi_u - 1)(\phi_v - 1) + \mathrm{e}^\theta - 1 \}} - \theta^{-2} \mathcal{C}\left(u, v; \theta\right) & \text{(Frank copula)} \end{cases},$$

$\phi_x = \log(-\log(x))$ in the Gumbel copula and $\phi_x = \mathrm{e}^{\theta x}$ in the Frank copula. Hence, we obtain the NR algorithm such that

$$\theta_l = \theta_{l-1} - \left\{ \frac{d}{d\theta} Q(\theta) \right\}^{-1} Q(\theta) \Big|_{\theta = \theta_{l-1}}, \quad l = 1, 2, 3, \ldots,$$

where θ_0 is the starting point of the algorithm.

One remark is that the numerical integrations is usually required to calculate the double integrals, such as (A.1) and (A.2). In the remainder of this paragraph, we discuss the numerical integration methods and their performances to calculate (A.1) and (A.2). Let $t_j = jh$, $j = 1, 2, 3, \ldots$ with $t_0 = 0$ and $h = 10/M$, and let $u_0 < u_1 < u_2 < \cdots < u_M$ be a partition of interval [0,1] with $u_0 = 0$ and $u_M = 1$. The integrals (A.1) and (A.2) can be approximated by

$$\begin{aligned} &\tilde{\rho}^{(1)}_{\tilde{M}}(\theta) = \textstyle\sum_{m=1}^{\tilde{M}} \sum_{l=1}^{\tilde{M}} \bar{\bar{\mathcal{C}}} \left(\mathrm{e}^{-t_m}, \mathrm{e}^{-t_l}; \theta\right) \delta t_m \delta t_l - 1 \\ \text{and} \quad &\tilde{\rho}^{(2)}_{M}(\theta) = \textstyle\sum_{m=1}^{M} \sum_{l=1}^{M} \bar{\bar{\mathcal{C}}} \left(u_m, u_l; \theta\right) \bar{u}_m^{-1} \bar{u}_l^{-1} \delta u_l \delta u_m - 1, \end{aligned}$$

respectively, where the notations of $\delta t_m = t_m - t_{m-1}$ and $\delta u_m = u_m - u_{m-1}$ are commonly used, in a trapezoidal rule

$$\begin{aligned} &\bar{\bar{\mathcal{C}}}\left(u_m, u_l; \theta\right) = \frac{1}{4}\left\{\mathcal{C}\left(u_m, u_l; \theta\right) + \mathcal{C}\left(u_m, u_{l-1}; \theta\right) + \mathcal{C}\left(u_{m-1}, u_l; \theta\right) + \mathcal{C}\left(u_{m-1}, u_{l-1}; \theta\right)\right\} \\ &\text{and} \quad \bar{u}_m^{-1} = (u_m^{-1} + u_{m-1}^{-1})/2 \end{aligned}$$

are used, in Simpson's rule

$$\begin{aligned} \bar{\bar{\mathcal{C}}}\left(u_m, u_l; \theta\right) = \frac{1}{36}\Big[&\mathcal{C}\left(u_m, u_l; \theta\right) + \mathcal{C}\left(u_m, u_{l-1}; \theta\right) + \mathcal{C}\left(u_{m-1}, u_l; \theta\right) + \mathcal{C}\left(u_{m-1}, u_{l-1}; \theta\right) \\ &+4\left\{\mathcal{C}\left(\tilde{u}_m, u_l; \theta\right) + \mathcal{C}\left(u_m, \tilde{u}_l; \theta\right) + \mathcal{C}\left(\tilde{u}_m, u_{l-1}; \theta\right) + \mathcal{C}\left(u_{m-1}, \tilde{u}_l; \theta\right)\right\} + 16\mathcal{C}\left(\tilde{u}_m, \tilde{u}_l; \theta\right)\Big], \\ \bar{u}_m^{-1} = \{u_m^{-1} + 4\tilde{u}_m^{-1} + u_{m-1}^{-1})/6 \quad &\text{and} \quad \tilde{u}_m = (u_m + u_{m-1})/2 \end{aligned}$$

are used. Also, $\tilde{M}$ for $\tilde{\rho}^{(1)}_{\tilde{M}}(\theta)$ is selected as the first integer satisfying

$$|\tilde{\rho}^{(1)}_{\tilde{M}+j}(\theta) - \tilde{\rho}^{(1)}_{\tilde{M}+j-1}(\theta)|/\tilde{\rho}^{(1)}_{\tilde{M}}(\theta) < 10^{-18}, \;\; j = 1, 2, 3,$$

which makes us escape from the loop computation.

Now, we see the performances of these numerical integrations. Table A.1 provides the values of $\tilde{\rho}^{(1)}_{\tilde{M}}(\theta)$ and $\tilde{\rho}^{(2)}_{M}(\theta)$ computed by the above two rules when M = 100, 200, 400, 800 and 1600 and θ's for the three copulas are given. The columns T* and S* (time) of Table A.1 are the computation values and times (second) using the trapezoidal and Simpson's rules, respectively. Computation times of Simpson's rules until reaching the objective ρ are shorter than trapezoidal rules. In particular, it is more efficient to compute $\tilde{\rho}^{(1)}_{\tilde{M}}(\theta)$ than $\tilde{\rho}^{(2)}_{M}(\theta)$. Hence, the algorithm to compute $\tilde{\rho}^{(1)}_{\tilde{M}}(\theta)$ by Simpson's rules is adopted in our computational program of Section E.

A.3 Generation of bivariate exponential survival data

We describe the generation of the random number of bivariate survival data. This simulation technique is necessary to compute empirical powers in the present article. Suppose that the marginals of T_{i1} and T_{i2} are exponential, that is, the marginal survival functions are $S_j^{(g_i)}(t) = \exp(-\lambda_j^{(g_i)} t)$, $j = 1, 2$. In the below, u_1 and u_2 are mutually independent random numbers from a standard uniform distribution $\mathrm{U}(0, 1)$.

Clayton copula. For $g_i = k$, (T^*_{i1}, T^*_{i2}) can be generated from

$$\begin{cases} T^*_{i1} = -\log(1 - u_1)/\lambda_1^{(k)} \\ T^*_{i2} = \log\left(1 - \exp\left(-\lambda_1^{(k)} T^*_{i1}\right)^{-\theta^{(k)}} + \exp\left(-\lambda_1^{(k)} T^*_{i1}\right)^{-\theta^{(k)}} (1 - u_2)^{-\theta^{(k)}/(1+\theta^{(k)})}\right) / \lambda_2^{(k)} \theta^{(k)} \end{cases}$$

using u_1 and u_2 (Prentice and Cai, 1992).

Table A.1: Performance of four numerical integrations calculated ρ given θ.

			$\tilde{\rho}^{(1)}_{\tilde{M}}(\theta)$		$\tilde{\rho}^{(2)}_{M}(\theta)$	
copula	θ	M	T*(time)	S*(time)	T*(time)	S*(time)
		100	0.433 (0.02)	0.300 (0.17)	0.289 (0.05)	0.294 (0.10)
		200	0.332 (0.08)	0.300 (0.55)	0.294 (0.20)	0.297 (0.44)
Clayton	0.3277	400	0.308 (0.26)	0.300 (2.17)	0.297 (0.77)	0.298 (1.73)
		800	0.302 (0.23)	0.300 (2.06)	0.298 (2.66)	0.299 (1.65)
		1600	0.301 (0.91)	0.300 (7.76)	0.299 (2.95)	0.300 (6.63)
		100	0.612 (0.01)	0.500 (0.04)	0.484 (0.01)	0.491 (0.02)
		200	0.526 (0.01)	0.500 (0.14)	0.491 (0.05)	0.495 (0.11)
	0.6415	400	0.506 (0.07)	0.500 (0.53)	0.496 (0.18)	0.498 (0.42)
		800	0.502 (0.24)	0.500 (2.01)	0.498 (0.75)	0.499 (1.65)
		1600	0.500 (0.91)	0.500 (7.76)	0.499 (2.95)	0.500 (6.63)
		100	0.877 (0.01)	0.800 (0.04)	0.784 (0.01)	0.791 (0.02)
		200	0.816 (0.02)	0.800 (0.13)	0.792 (0.05)	0.796 (0.11)
	1.7353	400	0.804 (0.08)	0.800 (0.53)	0.796 (0.17)	0.798 (0.41)
		800	0.801 (0.22)	0.800 (2.03)	0.798 (0.73)	0.799 (1.67)
		1600	0.800 (0.91)	0.800 (7.67)	0.799 (2.96)	0.800 (6.59)
		100	0.394 (0.00)	0.299 (0.04)	0.298 (0.02)	0.299 (0.04)
		200	0.317 (0.01)	0.300 (0.15)	0.299 (0.07)	0.300 (0.16)
Gumbel	0.7249	400	0.303 (0.07)	0.300 (0.56)	0.300 (0.30)	0.300 (0.63)
		800	0.301 (0.24)	0.300 (2.10)	0.300 (1.13)	0.300 (2.57)
		1600	0.300 (0.94)	0.300 (7.97)	0.300 (4.59)	0.300 (10.3)
		100	0.563 (0.01)	0.499 (0.04)	0.496 (0.02)	0.498 (0.04)
		200	0.509 (0.02)	0.500 (0.17)	0.498 (0.08)	0.499 (0.15)
	0.5582	400	0.501 (0.07)	0.500 (0.61)	0.499 (0.28)	0.500 (0.66)
		800	0.500 (0.29)	0.500 (2.35)	0.500 (1.16)	0.500 (2.55)
		1600	0.500 (1.02)	0.500 (8.93)	0.500 (4.62)	0.500 (10.3)
		100	0.845 (0.00)	0.800 (0.05)	0.790 (0.01)	0.795 (0.04)
		200	0.805 (0.02)	0.800 (0.19)	0.795 (0.07)	0.797 (0.14)
	0.3027	400	0.801 (0.08)	0.800 (0.68)	0.798 (0.29)	0.799 (0.64)
		800	0.800 (0.31)	0.800 (2.59)	0.799 (1.15)	0.799 (2.57)
		1600	0.800 (1.15)	0.800 (9.84)	0.799 (4.55)	0.800 (10.2)
		100	0.400 (0.04)	0.300 (0.01)	0.300 (0.02)	0.300 (0.03)
		200	0.322 (0.00)	0.300 (0.07)	0.300 (0.06)	0.300 (0.15)
Frank	-2.4882	400	0.305 (0.02)	0.300 (0.26)	0.300 (0.24)	0.300 (0.53)
		800	0.301 (0.10)	0.300 (0.91)	0.300 (0.99)	0.300 (2.20)
		1600	0.300 (0.39)	0.300 (3.59)	0.300 (3.97)	0.300 (8.87)
		100	0.568 (0.04)	0.500 (0.01)	0.500 (0.01)	0.500 (0.03)
		200	0.513 (0.01)	0.500 (0.06)	0.500 (0.06)	0.500 (0.13)
	-4.7299	400	0.503 (0.04)	0.500 (0.23)	0.500 (0.22)	0.500 (0.49)
		800	0.501 (0.10)	0.500 (0.87)	0.500 (0.90)	0.500 (2.04)
		1600	0.500 (0.39)	0.500 (3.42)	0.500 (3.62)	0.500 (8.12)
		100	0.850 (0.29)	0.800 (3.63)	0.799 (0.03)	0.800 (0.06)
		200	0.807 (0.63)	0.800 (6.34)	0.800 (0.11)	0.800 (0.23)
	-13.943	400	0.801 (0.51)	0.800 (7.91)	0.800 (0.42)	0.800 (0.92)
		800	0.800 (0.40)	0.800 (10.6)	0.800 (1.66)	0.800 (3.72)
		1600	0.800 (0.85)	0.800 (17.6)	0.800 (6.62)	0.800 (14.8)

Gumbel copula. When $g_i = k$, (T^*_{i1}, T^*_{i2}) following the Gumbel copula can be generated as

$$\begin{cases} T^*_{i1} = (1-u_1)^{\theta^{(k)}} W/\lambda_1^{(k)} \\ T^*_{i2} = \left\{ W^{1/\theta^{(k)}} - \left(\lambda_1^{(k)} T^*_{i1}\right)^{1/\theta^{(k)}} \right\}^{\theta^{(k)}} \Big/ \lambda_2^{(k)} \end{cases}$$

(Oakes and Manatunga, 1992), where W is the solution of the equation $(1 + W\theta^{(k)})\exp(-W) = 1 - u_2$. We can usually obtain the W by a simple Newton-Raphson algorithm.
Frank copula. When $g_i = k$, (T^*_{1i}, T^*_{2i}) from the Frank copula can be generated as

$$\begin{cases} T^*_{i1} = -\log(1-u_1)/\lambda_1^{(k)} \\ T^*_{i2} = -\log\left[\log\left[W/\left\{W + (1-\mathrm{e}^{\theta^{(k)}})(1-u_2)\right\}\right] \Big/ \theta^{(k)}\right] \Big/ \lambda_2^{(k)} \end{cases}$$

(Genest, 1987), where $W = \mathrm{e}^{\theta^{(k)}(1-u_1)} + \left\{\mathrm{e}^{\theta^{(k)}} - \mathrm{e}^{\theta^{(k)}(1-u_1)}\right\}(1-u_2)$.

B Technical details for Section 2.3

In Section 2.3 of the main text, the power formula (2.4) is constructed based on the asymptotic result of the bivariate logrank statistic. Here, we provide this asymptotic property and its proof.

B.1 Asymptotic distribution of bivariate logrank statistic

In advance, we express the bivariate weighted logrank statistic $(U_1, U_2)'$ using the counting and at-risk processes. Let $\mathcal{N}_{ij}(t) = \mathbb{1}(T_{ij} \le t, \Delta_{ij} = 1)$ and $\mathcal{Y}_{ij}(t) = \mathbb{1}(T_{ij} \ge t)$ be the counting and at-risk processes for the j^{th} endpoint of the i^{th} participant. Denote the total sums of $\mathcal{N}_{ij}$ and $\mathcal{Y}_{ij}$ in the group k by

$$\mathcal{N}_j^{(k)}(t) = \textstyle\sum_{i=1}^n \mathbb{1}_i^{(k)} \mathcal{N}_{ij}(t), \text{ and } \mathcal{Y}_j^{(k)}(t) = \textstyle\sum_{i=1}^n \mathbb{1}_i^{(k)} \mathcal{Y}_{ij}(t),$$

where $\mathbb{1}_i^{(k)} = \mathbb{1}(g_i = k)$. Then, the Nelson-Aalen estimator of $\Lambda_j^{(k)}(t)$ is written as

$$\widehat{\Lambda}_j^{(k)}(t) = \int_0^t d\mathcal{N}_j^{(k)}(s)/\mathcal{Y}_j^{(k)}(s).$$

The function $\widehat{H}_j(t)$ included in the weighted logrank statistic is usually a negative predictable function of bounded variation. Let $\mathcal{Y}_j(t) = \mathcal{Y}_j^{(1)}(t) + \mathcal{Y}_j^{(2)}(t)$ and $\widehat{H}_j^L(t) = n^{-1}\mathcal{Y}_j^{(1)}(t)\mathcal{Y}_j^{(2)}(t)/\mathcal{Y}_j(t)$. As well-known, the logrank statistic uses $\widehat{H}_j(t) = \widehat{H}_j^L(t)$, the Gehan-Wilcoxon statistic uses $\widehat{H}_j(t) = n^{-1}\mathcal{Y}_j(t)\widehat{H}_j^L(t)$ and the Prentice-Wilcoxon statistic selects $\widehat{H}_j(t) = n^{-1}\mathcal{Y}_j(t)\widehat{S}_{j\mathrm{p}}(t_-)\widehat{H}_j^L(t)$, where $\widehat{S}_{j\mathrm{p}}(t)$ is the Kaplan-Meier estimator for the j^{th} endpoint in the pooled sample.

For differentials of the bivariate function, define the notations

$$S^{(k)}(t, ds) = S^{(k)}(t, s) - S^{(k)}(t, s_-), \quad S^{(k)}(dt, s) = S^{(k)}(t, s) - S^{(k)}(t_-, s),$$

and $S^{(k)}(dt, ds) = S^{(k)}(t, ds) - S^{(k)}(t_-, ds) = S^{(k)}(dt, s) - S^{(k)}(dt, s_-)$. These mean

$$S^{(k)}(dt, ds) = \{\tfrac{\partial}{\partial s} S^{(k)}(dt, s)\}ds \text{ and } S^{(k)}(t, ds) = \{\tfrac{\partial}{\partial s} S^{(k)}(t, s)\}ds$$

if s is a continuity point of $S^{(k)}(dt, s)$ and $S^{(k)}(t, s)$. Similarly, we have $S^{(k)}(dt, s) = \{\frac{\partial}{\partial t} S^{(k)}(t, s)\}dt$ if t is its continuity point. See Prentice and Cai (1992) for further details of notations. We can give the following result (Theorem 1) on the asymptotic distribution of the bivariate weighted logrank statistic using these notations. In the main text, $a^{(k)}$ is given as the ratio of participants assigned to the group k to the total number n, but it is exactly as the limit of $a_n^{(k)} = n^{(k)}/n$ in this supplemental material, where $n^{(k)}$ is the numbers of participants in the group k.

Theorem 1. Let $y_j^{(k)}(s) = \mathrm{E}[\mathcal{Y}_{ij}^{(g_i)}(s)|g_i = k]$, $\tau = \inf\{t : y_j^{(1)}(t) \wedge y_j^{(2)}(t) = 0\}$, $a_n^{(k)} = n^{(k)}/n$ and $a^{(k)} = \lim_{n\to\infty} a_n^{(k)}$, where $t \wedge s = \min(t, s)$. Suppose that $\widehat{H}_j(t)$ converges in probability to $H_j(t)$ uniformly on $t \in [0, \tau]$ as $n \to \infty$, where $H_j(t)$ is a deterministic function of bounded variation. Then, for sufficiently large n, the distribution of $(U_1(\tau), U_2(\tau))'$ is approximately bivariate normal with mean vector $\sqrt{n}(\mu_1(\tau), \mu_2(\tau))'$ and variance-covariance matrix

$$\boldsymbol{V}(\tau) = \begin{pmatrix} V_{11}(\tau) & V_{12}(\tau) \\ V_{12}(\tau) & V_{22}(\tau) \end{pmatrix},$$

where

$$\sqrt{n}\mu_j(\tau) = \sqrt{n}\int_0^\tau H_j(t)\{d\Lambda_j^{(2)}(t) - d\Lambda_j^{(1)}(t)\}, \tag{B.1}$$

$$V_{jj}(\tau) = \int_0^\tau H_j(t)^2 \left\{ \frac{(1 - d\Lambda_j^{(1)}(t))d\Lambda_j^{(1)}(t)}{a^{(1)}y_j^{(1)}(t)} + \frac{(1 - d\Lambda_j^{(2)}(t))d\Lambda_j^{(2)}(t)}{a^{(2)}y_j^{(2)}(t)} \right\}, \tag{B.2}$$

$$V_{12}(\tau) = \int_0^\tau\int_0^\tau H_1(t)H_2(s)C(t,s) \left\{ \frac{dA^{(1)}(t,s)}{a^{(1)}y_1^{(1)}(t)y_2^{(1)}(s)} + \frac{dA^{(2)}(t,s)}{a^{(2)}y_1^{(2)}(t)y_2^{(2)}(s)} \right\}, \tag{B.3}$$

$$\begin{aligned} dA^{(k)}(t,s) = S^{(k)}(dt,ds) + S^{(k)}(t_-,ds)d\Lambda_1^{(k)}(t) + S^{(k)}(dt,s_-)d\Lambda_2^{(k)}(s) \\ +S^{(k)}(t_-,s_-)d\Lambda_1^{(k)}(t)d\Lambda_2^{(k)}(s) \end{aligned}$$

and $C(t, s)$ is the joint survival function $\Pr(C_{i1} > t, C_{i2} > s)$ for censoring variables.

B.2 The derivation for the power formula (2.4)

The conditional variance of $U_j(t)$, based on the hypergeometric distribution theory under $\mathcal{H}_0$, is

$$\widehat{V}_{jj}^0(t) = n\int_0^t \widehat{H}_j(s)^2 \frac{\mathcal{Y}_j(s)}{\mathcal{Y}_j^{(1)}(s)\mathcal{Y}_j^{(2)}(s)} \left\{1 - \frac{d\mathcal{N}_j(s) - 1}{\mathcal{Y}_j(s) - 1}\right\} \frac{d\mathcal{N}_j(s)}{\mathcal{Y}_j(s)},$$

where $\mathcal{N}_j(s) = \mathcal{N}_j^{(1)}(s) + \mathcal{N}_j^{(2)}(s)$. Similarly to discussion for Theorem 1, the limit form of $\widehat{V}_{jj}^0(\tau)$ is found out as

$$V_{jj}^0(\tau) = \int_0^\tau H_j(t)^2 \left\{ \frac{a^{(1)}y_j^{(1)}(t) + a^{(2)}y_j^{(2)}(t)}{a^{(1)}a^{(2)}y_j^{(1)}(t)y_j^{(2)}(t)} \right\} (1 - d\Lambda_j^y(t))d\Lambda_j^y(t), \tag{B.4}$$

where

$$d\Lambda_j^y(t) = \frac{a^{(1)}y_j^{(1)}(t)\Lambda_j^{(1)}(dt) + a^{(2)}y_j^{(2)}(t)\Lambda_j^{(2)}(dt)}{a^{(1)}y_j^{(1)}(t) + a^{(2)}y_j^{(2)}(t)}.$$

In order to derive a simple power formula in the main text, we consider the test statistic $Z_j^0 = -U_j(\tau)/\sqrt{V_{jj}^0(\tau)}$ rather than $\widehat{Z}_j^0 = -U_j(\tau)/\sqrt{\widehat{V}_{jj}^0(\tau)}$, where Z_j^0 is $\widehat{Z}_j^0$ of which the denominator

$\widehat{V}_{jj}^0(\tau)$ is replaced by the limit form $V_{jj}^0(\tau)$ $(j=1,2)$. That is, the testing procedure (2.3) of the weighted logrank statistic is modified such that

$$\text{reject } \mathcal{H}_0 \text{ if and only if } Z_1^0 > \mathrm{z}_\alpha \text{ and } Z_2^0 > \mathrm{z}_\alpha. \tag{B.5}$$

Hence, the power for the testing procedure (2.3) is approximately obtained as (2.4) using the rejection region (B.5) and the asymptotic distribution of $(Z_1^0, Z_2^0)'$, because, as written in the main text, Theorem 1 provides that $(Z_1^0, Z_2^0)'$ is approximately bivariate normally distributed with mean vector $\sqrt{n}\boldsymbol{\delta}$ and variance-covariance matrix $\boldsymbol{\Sigma}$.

B.3 Proof of Theorem 1

Here we provide a proof of Theorem 1 in the previous section.

Let $\mathcal{M}_{ij}(t) = \mathcal{N}_{ij}(t) - \int_0^t \mathcal{Y}_{ij}(s) d\Lambda_j^{(g_i)}(s)$ and $\mathscr{F}_{j,t} = \sigma\{\mathcal{N}_{ij}(s), \mathcal{N}_{ij}^C(s) : 0 \le s \le t, i = 1, \cdots, n\}$, where $\mathcal{N}_{ij}^C(t) = \mathbb{1}(T_{ij} \le t, \Delta_{ij} = 0)$ is a counting process for the censoring time. That is, $\{\mathscr{F}_{j,t} : t \ge 0\}$ is a standard filtration generated from the history through time t for the j^{th} endpoints, so that $\mathcal{M}_{ij}(t)$ has the $\mathscr{F}_{j,t}$-martingale property. By the definition of U_j, we have $U_j(t) = \sqrt{n}\widehat{\mu}_j(t) + U_j^{\mathcal{M}}(t)$, where

$$\sqrt{n}\widehat{\mu}_j(t) = \int_0^t \widehat{H}_j(s)\{d\Lambda_j^{(2)}(s) - d\Lambda_j^{(1)}(s)\},$$

$$U_j^{\mathcal{M}}(t) = \sqrt{n}\sum_{i=1}^n \int_0^t \widehat{H}_j(s)\left\{\frac{\mathbb{1}_i^{(2)} d\mathcal{M}_{ij}(s)}{\mathcal{Y}_j^{(2)}(s)} - \frac{\mathbb{1}_i^{(1)} d\mathcal{M}_{ij}(s)}{\mathcal{Y}_j^{(1)}(s)}\right\}.$$

Using well-known martingale theory for survival analysis (Fleming and Harrington, 1991; Andersen *and others*, 1993), as $n \to \infty$, the $\mathscr{F}_{j,t}$-martingale process $U_j^{\mathcal{M}}(t)$ converges in distribution to the Gaussian process with the predictable variance process

$$\begin{aligned}\langle U_j^{\mathcal{M}}\rangle(t) &= n\int_0^t \widehat{H}_j(s)^2 \left\{\frac{(1-\Lambda_j^{(1)}(ds))\Lambda_j^{(1)}(ds)}{\mathcal{Y}_j^{(1)}(s)} + \frac{(1-\Lambda_j^{(2)}(ds))\Lambda_j^{(2)}(ds)}{\mathcal{Y}_j^{(2)}(s)}\right\} \\ &\overset{\text{a.s.}}{\to} \text{(B.2)} \quad (n\to\infty),\end{aligned}$$

where $\overset{\text{a.s.}}{\to}$ denotes almost sure convergence. Also, as $n \to \infty$, $\mathcal{Y}_j^{(k)}(s)/n^{(k)}$ converges almost surely to $y_j^{(k)}(s)$ uniformly on $s \in [0,\tau]$ by the Glivenko-Cantelli theorem.

To obtain the covariance of $U_1^{\mathcal{M}}$ and $U_2^{\mathcal{M}}$, we need $\mathrm{E}\left[d\mathcal{M}_{i1}(t)d\mathcal{M}_{i2}(s)\right]$. This is formulated as

$$\begin{aligned}\mathrm{E}\left[d\mathcal{M}_{i1}(t)d\mathcal{M}_{i2}(s)\right] &= \mathrm{E}\left[\mathcal{Y}_{i1}(t)\mathcal{Y}_{i2}(s)\mathrm{E}\left[d\mathcal{M}_{i1}(t)d\mathcal{M}_{i2}(s)|\mathcal{Y}_{i1}(t)\mathcal{Y}_{i2}(s)\right]\right] \\ &= \Pr(\mathcal{Y}_{i1}(x)\mathcal{Y}_{i2}(y)=1)\mathrm{E}\left[d\mathcal{M}_{i1}(x)d\mathcal{M}_{i2}(y)|\mathcal{Y}_{i1}(x)\mathcal{Y}_{i2}(y)=1\right] \\ &= \begin{cases} C(t_-,s_-)S^{(1)}(t_-,s_-)dB^{(1)}(t,s) & \text{if } g_i = 1, \\ C(t_-,s_-)S^{(2)}(t_-,s_-)dB^{(2)}(t,s) & \text{if } g_i = 2,\end{cases}\end{aligned}$$

(see Prentice and Cai (1992) and Jung (2008)), where $dB^{(k)}(t,s)$ can be expressed as

$$dB^{(k)}(t,s) = \Lambda^{(k)}(dt,ds) - \Lambda^{(k)}(dt,s)\Lambda_2^{(k)}(ds) - \Lambda^{(k)}(t,ds)\Lambda_1^{(k)}(dt) + \Lambda_1^{(k)}(dt)\Lambda_2^{(k)}(ds)$$

and $\Lambda^{(g_i)}(t,s)$ is the joint cumulative hazard function of (T_{i1}, T_{i2}), so that we have

$$dA^{(k)}(t,s) = S^{(k)}(t_-,s_-)dB^{(k)}(t,s).$$

On the covariance of $U_1^{\mathcal{M}}$ and $U_2^{\mathcal{M}}$ from these results, we obtain

$$\begin{aligned}
&\mathrm{E}[U_1^{\mathcal{M}}(t)U_2^{\mathcal{M}}(s)] \\
&= n\sum_{i=1}^{n}\int_0^t\int_0^s \mathrm{E}\left[\widehat{H}_1(x)\widehat{H}_2(y)\left\{\frac{\mathbb{1}_i^{(1)}d\mathcal{M}_{i1}(x)d\mathcal{M}_{i2}(y)}{\mathcal{Y}_1^{(1)}(x)\mathcal{Y}_2^{(1)}(y)}+\frac{\mathbb{1}_i^{(2)}d\mathcal{M}_{i1}(x)d\mathcal{M}_{i2}(y)}{\mathcal{Y}_1^{(2)}(x)\mathcal{Y}_2^{(2)}(y)}\right\}\right] \\
&\overset{\text{a.s.}}{\to} \ (\text{B.3}) \quad (n\to\infty),
\end{aligned}$$

which can be shown using the $\mathscr{F}_{j,t}$-martingale property of $\mathcal{M}_{ij}(t)$ and the uniform convergences in probability of $\mathcal{Y}_j^{(k)}(\cdot)$ and $\widehat{H}_j(\cdot)$.

Let $J(x,\mathrm{s})=\sup\{|x(t)-x(t_-)| : \mathrm{s}<t\le\tau\}$, which is the absolute value of the maximum jump in x over the interval $[\mathrm{s},\tau]$. If both T_{i1} and T_{i2} are continuous random variables, letting $\mathrm{s}=0$ and then we have

$$\lim_{n\to\infty}\mathrm{E}[J(\langle U_j^{\mathcal{M}},U_l^{\mathcal{M}}\rangle,0)]=0 \quad\text{and}\quad \lim_{n\to\infty}\left\{\mathrm{E}[J(U_j^{\mathcal{M}},0)^2]\right\}=0,$$

so that $(U_1^{\mathcal{M}}(\cdot),U_2^{\mathcal{M}}(\cdot))'$ is asymptotically distributed as a bivariate Gaussian martingale process by Whitt (2007, Theorem 2.1). While, if either of T_{i1} and T_{i2}, for example, T_{i2} is discrete random variable, considering the case of $\mathrm{s}=\tau$, we have a result restricted to one time point of Whitt (2007, Theorem 2.1), so that we can show that the asymptotic distribution of $(U_1^{\mathcal{M}}(\tau),U_2^{\mathcal{M}}(\tau))'$ is bivariate normal. In addition, it is immediate from the condition that as $n\to\infty$, $\widehat{\mu}_j(t)$ converges in probability to $\mu_j(t)$ uniformly on $[0,\tau]$. Therefore, a proof of this theorem is complete. □

C Technical details for Section 3

C.1 The selection of the logrank statistic

The main text limits to the logrank statistic for testing (2.1) on and after Section 3, so that the limit function of $\widehat{H}_j(t)$ is

$$H_j(t)=H_j^L(t)=\frac{a^{(1)}a^{(2)}y_j^{(1)}(t)y_j^{(2)}(t)}{a^{(1)}y_j^{(1)}(t)+a^{(2)}y_j^{(2)}(t)},$$

which is obtained following $\mathrm{E}[\mathcal{Y}_{ij}(t)]=a_n^{(1)}y_j^{(1)}(t)+a_n^{(2)}y_j^{(2)}(t)$ and $y_j^{(k)}(t)=S_j^{(k)}(t_-)C_j(t_-)$, where $C_j(t)=\Pr(t<C_{ij})$ is the marginal survival function of C_{ij}. In addition, assuming that censoring times are the same, i.e., $C_1(t)=C_2(t)$, we have the form of

$$H_j(t)=a^{(1)}a^{(2)}\frac{C(t_-)S_j^{(1)}(t_-)S_j^{(2)}(t_-)}{S_j^{\mathrm{p}}(t_-)} \quad\text{and}\quad C(t,s)=C(t\vee s),$$

where $S_j^{\mathrm{p}}(t)=a^{(1)}S_j^{(1)}(t)+a^{(2)}S_j^{(2)}(t)$ and $t\vee s=\max(t,s)$. This form of $H_j(t)$ is used in Section 3 of the main text.

C.2 Computational efficiency of the numerical integration (Section 3.1)

We investigate computational efficiency of numerical integration for the bivariate logrank statistic discussed in Section 3.1 of the main text. Generally Simpson's rule has higher efficiency than a trapezoidal rule given in (3.1), while the original expression of Simpson's rule is quite complicated. The modification of (3.1) based on Simpson's rule is

$$
\begin{aligned}
\bar{G}(t_m) &= \left\{G(t_m) + 4G(\tilde{t}_m) + G(t_{m-1})\right\}/6, \\
\bar{\bar{G}}(t_m, s_l) &= \frac{1}{36}\Big[G(t_m, t_l) + G(t_m, t_{l-1}) + G(t_{m-1}, t_l) + G(t_{m-1}, t_{l-1}) \\
&\quad +4\left\{G(\tilde{t}_m, t_l) + G(t_m, \tilde{t}_l) + G(\tilde{t}_m, t_{l-1}) + G(t_{m-1}, \tilde{t}_l)\right\} \\
&\quad +16G(\tilde{t}_m, \tilde{t}_l)\Big] \text{ and } \tilde{t}_m = (t_m + t_{m-1})/2
\end{aligned}
\tag{C.1}
$$

similarly to the form in Section A.2. There are trade-offs between the expression, precision and cost. Hence, here we compare the trapezoidal method of the main text with two Simpson's methods (one is simple replacement and another is simultaneous application).

Table C.1 shows the values of the δ_j and $\sigma_{12}/\sigma_1\sigma_2$ (see Section 3.2 of the main text on the notions) computed by trapezoidal method (T) written in the main text, its simple replacement (S_1) to Simpson's rule and simultaneous application (S_2) of Simpson's rule to the overall integrand, where $\rho^{(1)} = \rho^{(2)} = 0.8$, $S_1^{(1)}(\tau) = S_1^{(2)}(\tau) = 0.1$, $\tau_{\mathrm{a}} = 2$, $\tau_{\mathrm{f}} = 3$, $\psi_1^{-1} = \psi_2^{-1} = 1.2, 1.5$ and $M = 100 \times 2^{l-1}$, $l = 1, \ldots, 5$ are used. The column "time" of Table C.1 indicates times until the methods T and S_1 compute all of the bivariate logrank statistic. We consider the results of the method S_2 with the largest M as a standard, because what applies Simpson's rule simultaneously to the overall integrand is the usual Simpson's method in numerical integration. We find that the method T is not so inferior to the method S_2 even if M is small, such as 100, and the method S_1 has the almost same results as S_2. Hence, for a practical use, the method S_1 with $M = 100$ is adopted as default in our computational program of Section E.

C.3 The derivation for sample size formula (3.2) (Section 3.2)

We show how the sample size formula (3.2) is derived from the power formula (2.4) in the main text. Considering a transformed bivariate logrank statistic, $(Z_1^0\sigma_1^0/\sigma_1, Z_2^0\sigma_2^0/\sigma_2)'$ and a symmetry property and shift to $-\sqrt{n}\boldsymbol{\delta}$ on the integration (2.4), where $\boldsymbol{\delta} = (\delta_1, \delta_2)'$. When n is sufficiently large, Theorem 1 states that Z_j can be approximated by normal distribution with zero mean and variance σ_j^0/σ_j under the null hypothesis $\mathcal{H}_0$, $j = 1, 2$, while $(Z_1, Z_2)'$ can be approximated by bivariate normal distribution with mean vector $\sqrt{n}\boldsymbol{\delta}$ and variance-covariance matrix $\boldsymbol{R}$ given under true parameters. Hence, (2.4) can be transformed into

$$
1 - \beta \;=\; \int_{-\infty}^{\sqrt{n}\delta_1 - \mathrm{z}_\alpha\sigma_1^0/\sigma_1} \int_{-\infty}^{\sqrt{n}\delta_2 - \mathrm{z}_\alpha\sigma_2^0/\sigma_2} f(z_1, z_2; \boldsymbol{R}) dz_2 dz_1. \tag{C.2}
$$

This equation (C.2) is equivalent to the simultaneous equations

$$
1 - \beta = \int_{-\infty}^{K^{(1)}} \int_{-\infty}^{K^{(2)}} f(z_1, z_2; \boldsymbol{R}) dz_2 dz_1 \tag{C.3a}
$$

$$
\text{and} \qquad K^{(k)} = \sqrt{n}\delta_k - \mathrm{z}_\alpha\sigma_k^0/\sigma_k, \;\; k = 1, 2. \tag{C.3b}
$$

Because (C.3b) is equivalent to two linear equations

$$
n = \left(K^{(1)} + \frac{\sigma_1^0}{\sigma_1}\mathrm{z}_\alpha\right)^2 / \delta_1^2 \quad \text{and} \quad n = \left(K^{(2)} + \frac{\sigma_2^0}{\sigma_2}\mathrm{z}_\alpha\right)^2 / \delta_2^2,
$$

the formula (3.2) can be derived for the total sample size n by letting $K_\beta = K^{(2)}$. Given $K_\beta = K^{(2)}$, $K^{(1)}$ is expressed as

$$
K^{(1)} = \frac{\delta_1}{\delta_2}K_\beta + \mathrm{z}_\alpha\left(\frac{\sigma_2^0}{\sigma_2}\frac{\delta_1}{\delta_2} - \frac{\sigma_1^0}{\sigma_1}\right), \tag{C.4}
$$

Table C.1: The values of δ_j and $\sigma_{12}/\sigma_1\sigma_2$ computed from the trapezoidal (T) and two Simpson's (S_1, S_2) methods when $\rho^{(1)} = \rho^{(2)} = 0.8$, $S_1^{(1)}(\tau) = S_1^{(2)}(\tau) = 0.1$, $\tau_a = 2$, $\tau_f = 3$ and $\psi_1^{-1} = \psi_2^{-1}$

ψ_j^{-1}	copula	M	δ_j			$\sigma_{12}/\sigma_1\sigma_2$			time (s)	
			T	S_1	S_2	T	S_1	S_2	T	S_1
1.2	Clayton	100	-0.081496	-0.081495	-0.081495	0.695825	0.695931	0.695926	0.10	0.12
		200	-0.081495	-0.081495	-0.081495	0.695906	0.695932	0.695931	0.28	0.45
		400	-0.081495	-0.081495	-0.081495	0.695926	0.695932	0.695932	1.17	1.82
		800	-0.081495	-0.081495	-0.081495	0.695931	0.695933	0.695932	4.69	7.29
		1600	-0.081495	-0.081495	-0.081495	0.695932	0.695933	0.695933	18.8	28.8
	Gumbel	100	-0.081496	-0.081495	-0.081495	0.791356	0.791495	0.791489	0.11	0.15
		200	-0.081495	-0.081495	-0.081495	0.791460	0.791495	0.791493	0.44	0.67
		400	-0.081495	-0.081495	-0.081495	0.791486	0.791495	0.791495	1.75	2.76
		800	-0.081495	-0.081495	-0.081495	0.791493	0.791495	0.791495	7.08	11.0
		1600	-0.081495	-0.081495	-0.081495	0.791494	0.791495	0.791495	28.1	43.6
	Frank	100	-0.081496	-0.081495	-0.081495	0.863740	0.863877	0.863871	0.12	0.19
		200	-0.081495	-0.081495	-0.081495	0.863845	0.863879	0.863877	0.50	0.77
		400	-0.081495	-0.081495	-0.081495	0.863870	0.863879	0.863878	1.99	3.07
		800	-0.081495	-0.081495	-0.081495	0.863877	0.863879	0.863879	8.00	12.4
		1600	-0.081495	-0.081495	-0.081495	0.863878	0.863879	0.863879	31.9	49.6
1.5	Clayton	100	-0.173694	-0.173693	-0.173693	0.682905	0.683004	0.683001	0.08	0.12
		200	-0.173693	-0.173693	-0.173693	0.682981	0.683005	0.683004	0.28	0.44
		400	-0.173693	-0.173693	-0.173693	0.682999	0.683005	0.683005	1.16	1.80
		800	-0.173693	-0.173693	-0.173693	0.683004	0.683005	0.683005	4.65	7.24
		1600	-0.173693	-0.173693	-0.173693	0.683005	0.683005	0.683005	18.6	28.8
	Gumbel	100	-0.173694	-0.173693	-0.173693	0.786875	0.787000	0.786996	0.11	0.21
		200	-0.173693	-0.173693	-0.173693	0.786969	0.787000	0.786999	0.41	0.68
		400	-0.173693	-0.173693	-0.173693	0.786992	0.787000	0.787000	1.75	2.75
		800	-0.173693	-0.173693	-0.173693	0.786998	0.787000	0.787000	7.09	10.9
		1600	-0.173693	-0.173693	-0.173693	0.787000	0.787000	0.787000	28.2	43.8
	Frank	100	-0.173694	-0.173693	-0.173693	0.859370	0.859494	0.859490	0.12	0.19
		200	-0.173693	-0.173693	-0.173693	0.859465	0.859495	0.859495	0.50	0.76
		400	-0.173693	-0.173693	-0.173693	0.859488	0.859496	0.859496	1.97	3.13
		800	-0.173693	-0.173693	-0.173693	0.859494	0.859496	0.859496	7.95	12.3
		1600	-0.173693	-0.173693	-0.173693	0.859495	0.859496	0.859496	32.0	49.5

so that the integral equation (3.3) is obtained by substituting these equations into (C.3a).

Now, we provide a Newton-Raphson (NR) algorithm to solve the integral equation (3.3). Let

$$Q(K_\beta) = (1-\beta) - \int_{-\infty}^{\frac{\delta_1}{\delta_2}K_\beta + z_\alpha\left(\frac{\sigma_2^0}{\sigma_2}\frac{\delta_1}{\delta_2} - \frac{\sigma_1^0}{\sigma_1}\right)} \int_{-\infty}^{K_\beta} f\left(z_1, z_2; \boldsymbol{R}\right) dz_2 dz_1.$$

For the NR method, we need the first derivative to find $Q(K_\beta) = 0$. The first derivative is

$$\frac{dQ(K_\beta)}{dK_\beta} = -F_{\boldsymbol{R}}^{(1)}(K^{(1)}, K_\beta)\frac{\delta_1}{\delta_2} - F_{\boldsymbol{R}}^{(2)}(K^{(1)}, K_\beta),$$

where $K^{(1)}$ satisfies (C.4),

$$F_{\boldsymbol{R}}^{(1)}(K^{(1)}, K^{(2)}) = \int_{-\infty}^{K^{(2)}} f(K^{(1)}, z_2; \boldsymbol{R}) dz_2 \text{ and } F_{\boldsymbol{R}}^{(2)}(K^{(1)}, K^{(2)}) = \int_{-\infty}^{K^{(1)}} f(z_1, K^{(2)}; \boldsymbol{R}) dz_1.$$

Hence, we obtain the NR algorithm

$$K_\beta^{(l)} = K_\beta^{(l-1)} - \left\{\frac{dQ(K_\beta)}{dK_\beta}\right\}^{-1} Q(K_\beta)\bigg|_{K_\beta = K_\beta^{(l-1)}}, \quad l = 1, 2, \cdots,$$

where $K_\beta^{(0)}$ is the starting point of this algorithm. For example, as such a starting point, we may use z_β or $\mathrm{z}_{1-\sqrt{1-\beta}}$. In addition, we can compute $F_{\boldsymbol{R}}^{(k)}$ approximately from the numerical derivative using the CDF of the bivariate normal distribution. In Section E, we provide the R codes (R Development Core Team, 2005) included as the name `KBsolution` in the R function `samplesize.2sep` of Section E.

C.4 Technical details for Section 3.3

The design methods of the sample size in the article for the testing procedure (2.3) are classified to the following three cases:

Case 1. when censoring times are the same , i.e., $C_{i1} = C_{i2}$,
Case 2. special case of Case 1: a fixed-time censoring and proportional hazard alternatives,
Case 3. special case of Case 2: uncensored data and proportional hazard alternatives.

The sample size formula (3.2), which is derived under the logrank test, can be applied to Case 1 without assuming proportional hazard alternatives. So, the formula (3.2) under Case 3 reduces to the number of events given in Section 3.3 of the main text. However, Cases 2 and 3 include an important aspect to simplify a complicated situation in practice. We first discuss Case 2 in calculation for Σ and then Case 3 as further one.

Similarly to Section 3.1 of the main text, $\mu_j(\tau)$, $V_{jj}(\tau)$ and $V_{12}(\tau)$ can be approximated by numerical integration, which corresponds to providing $C(t) = 1$ on the time $t \in [0, \tau)$ (except τ) into $\mu_j^{[M]}$, $V_{jj}^{[M]}$ and $V_{12}^{[M]}$. Then, given an additional condition (a small effect size), a further simplification can be obtained. A simple computation in Case 2, under a proportional hazard alternative $\psi_j = \psi_j(t)$, is similar to Freedman (1982). Let $\gamma_j(t)$ be the ratio of the numbers at risk for the j^{th} endpoint in the two groups, $\widehat{\gamma}_j(t) = \mathcal{Y}_j^{(2)}(t)/\mathcal{Y}_j^{(1)}(t)$ and $\gamma_j(t) = a^{(2)}S_j^{(2)}(t_-)/a^{(1)}S_j^{(1)}(t_-)$. We can consider an approximation for the weighted function

$$\widehat{H}_j^L(t) = n^{-1}\mathcal{Y}_j^{(1)}(t)\widehat{\gamma}_j(t)/\left(1+\widehat{\gamma}_j(t)\right) \xrightarrow{\text{a.s.}} H_j^L(t) \overset{\psi_j \to 1}{\approx} 2a^{(1)}a^{(2)}S_j^{(1)}(t)/(1+\psi_j),$$

using $\gamma_j(t)/(1+\gamma_j(t)) \approx 2a^{(2)}/(1+\psi_j)$ when ψ_j is near 1. This gives an approximation of $\mu_j(\tau)$,

$$\mu_j(\tau) \overset{\psi_j \to 1}{\approx} 2a^{(1)}a^{(2)}\left((\psi_j-1)/(\psi_j+1)\right)\left(1-S_j^{(1)}(\tau)\right) = \mu_j^{\dagger}(\tau).$$

Similar argument yields approximations of $V_{jj}(\tau)$, $j = 1, 2$ and $V_{12}(\tau)$,

$$V_{jj}(\tau) \overset{\psi_j \to 1}{\approx} 4a^{(1)}a^{(2)}\left\{a^{(2)}\left(1-S_j^{(1)}(\tau)\right)/(1+\psi_j)^2 + a^{(1)}\left(1-S_j^{(2)}(\tau)\right)/(1+\bar{\psi}_j)^2\right\} = V_{jj}^{\dagger}(\tau),$$

$$V_{12}(\tau) \overset{\psi_j \to 1}{\approx} 4a^{(1)}a^{(2)}\left\{\frac{a^{(2)}A^{(1)}(\tau,\tau)}{(1+\psi_1)(1+\psi_2)} + \frac{a^{(1)}A^{(2)}(\tau,\tau)}{(1+\bar{\psi}_1)(1+\bar{\psi}_2)}\right\} = V_{12}^{\dagger}(\tau),$$

where $\bar{\psi}_j = 1/\psi_j$ and $A^{(k)}(t,s) = \int_0^t\int_0^s dA^{(k)}(x,y)$. Similarly, $V_{jj}^0(\tau)$ are approximated as

$$V_{jj}^0(\tau) \overset{\psi_j \to 1}{\approx} 4a^{(1)}a^{(2)}\left\{a^{(1)}\left(1-S_j^{(1)}(\tau)\right)/(1+\psi_j)^2 + a^{(2)}\left(1-S_j^{(2)}(\tau)\right)/(1+\bar{\psi}_j)^2\right\} = V_{jj}^{0\dagger}(\tau).$$

Note that $A^{(k)}(t,s)$ is $\mathrm{E}[\mathcal{M}_{i1}(t)\mathcal{M}_{i2}(s)|g_i = k]$ for the i^{th} participant belonging to the group k. Prentice and Cai (1992) gave one expression of $A^{(k)}(\tau,\tau)$,

$$\begin{aligned} A^{(k)}(\tau,\tau) \quad = \quad & S^{(k)}(\tau,\tau) - 1 + \textstyle\int_0^\tau S^{(k)}(t_-,\tau)d\Lambda_1^{(k)}(t) + \int_0^\tau S^{(k)}(\tau,s_-)d\Lambda_2^{(k)}(s) \\ & + \textstyle\int_0^\tau\int_0^\tau S^{(k)}(t_-,s_-)d\Lambda_1^{(k)}(t)d\Lambda_2^{(k)}(s). \end{aligned} \tag{C.5}$$

Let $A_1^{(k)}(\tau) = S^{(k)}(\tau,\tau) - 1$ (the 1st term), and let $A_2^{(k)}(\tau)$, $A_3^{(k)}(\tau)$ and $A_4^{(k)}(\tau)$ be the 2nd-4th terms in the right-hand side of (C.5). If $\mathcal{C}$ is of the Clayton copula, the 2nd and 3rd terms can be expressed by a hypergeometric function. However, it is generally necessary to perform the numerical integration for the 2nd-4th terms, because the explicit forms of the integrals for $A_2^{(k)}(\tau)$ and $A_3^{(k)}(\tau)$ are not obtained in the other copula models (Gumbel, Frank). In addition, $A_4^{(k)}(\tau)$ can not be expressed by well-known functions even if $\mathcal{C}$ is of Clayton's copula. Hence, the numerical integration method of Section 3 written under a general censoring scheme is still useful in Case 2.

Consider Case 3 (uncensored data) based on the above-mentioned. Since $S_j^{(k)}(\tau) = 0$, $j = 1, 2$ are imposed further (i.e., as $\tau \to \infty$), the approximations ($\mu_j^{\dagger}(\infty)$ and $V_{jj}^{\dagger}(\infty)$) of $\mu_j(\tau)$ and $V_{jj}(\tau)$ are more easily obtained from the results of Case 2. In particular, note that $A_2^{(k)}(\tau)$ and $A_3^{(k)}(\tau)$ are zeros when $S_1^{(k)}(\tau) = S_2^{(k)}(\tau) = 0$ (i.e., $A_2^{(k)}(\infty) = 0$ and $A_3^{(k)}(\infty) = 0$). Therefore, by (2.2), we have $A^{(k)}(\tau) = \rho^{(k)}$ under Case 3 (i.e., $A^{(k)}(\infty) = \rho^{(k)}$). For example, in Clayton's copula, the explicit value of $A_4^{(k)}(\tau) = \rho^{(k)} + 1$ (as $\tau \to \infty$) can be obtained for some special values of $\theta^{(k)}$, such as

$$A_4^{(k)}(\tau)\Big|_{S_1^{(k)}(\tau)\to 0, S_2^{(k)}(\tau)\to 0} = \begin{cases} 2(\pi^2 - 9)/3 \doteq 0.579 & \text{if } \theta^{(k)} = -0.5 \\ 1 & \text{if } \theta^{(k)} = 0 \\ \pi^2/6 \doteq 1.645 & \text{if } \theta^{(k)} = 1 \\ 2 & \text{if } \theta^{(k)} = \infty \end{cases} .$$

D Additional numerical experiments and their results

We here provide some numerical results omitted in Section 4.1 of the main text. These results are constructed based on a few minor changes about some settings (e.g., $S_j^{(1)}(\tau)$, ψ_j, $a^{(1)}$) given to create Table 2. Similarly to Section 4.1, the target power of $1 - \beta = 0.8$, the significance level of $\alpha = 0.025$, and the censoring distribution $C(t) = \mathbb{1}(t < \tau_{\mathrm{a}} + \tau_{\mathrm{f}}) - \mathbb{1}(\tau_{\mathrm{f}} \le t < \tau_{\mathrm{a}} + \tau_{\mathrm{f}})(t - \tau_{\mathrm{f}})/\tau_{\mathrm{a}}$ with $\tau_{\mathrm{a}} = 2$ and $\tau_{\mathrm{f}} = 3$ are used throughout. Also, the manners to create and present the numerical results here are used similarly to those in Table 2, such as the number of MC trials (100,000), exponential marginals of T_{i1} and T_{i2}, the levels of the correlations $\rho^{(k)}$ and the notations for the sample sizes n, n_{sim}, n_{ind} and n_{min} from (3.2) and three PSs and their empirical powers $\tilde{p}_{12}$, etc.

The cases of different baseline survival rates. In practice, the baseline survival rates are usually different by the endpoints (e.g., $S_1^{(1)}(\tau) \neq S_2^{(1)}(\tau)$), although the condition of equal τ-time survival rates ($S_1^{(1)}(\tau) = S_2^{(1)}(\tau)$) is assumed for simplicity in the experiment of Table 2 of the main text. We will see one example with different baseline survival rates by adopting the same settings (e.g., equal group size ratio $a^{(1)} = 0.5$, the combinations of hazard ratios, etc.) as Table 2 except for the condition of $S_1^{(1)}(\tau) = S_2^{(1)}(\tau)$. For comparison with the result of Table 2, set $(S_1^{(1)}(\tau), S_2^{(1)}(\tau)) = (0.1, 0.5)$ and $(0.5, 0.1)$ into the pairs of two τ-time survival rates. Table D.1 shows the result of the experiment conducted under these settings, where the column of $\boldsymbol{S}_{\tau}^{(1)}$ displays the values of $(S_1^{(1)}(\tau), S_2^{(1)}(\tau))$.

The tendencies of the sample sizes seen in Table D.1 are intermediate between the sizes corresponding to $S_j^{(1)}(\tau) = 0.1$ and 0.5 in Table 2, because one endpoint has fewer censored observations in compared with Table 2 while another endpoint is more censored. We find that, as compared with n_{sim}, n_{ind} and n_{min}, the sample size n from (3.2) has tendencies similar to ones found in Table 2. Hence, the formula (3.2) is also preferable in this situation in the sense that it provides slightly conservative results, which strengthens the claim of Section 4.1.

Table D.1: Total numbers of participants (n) calculated from (3.2) and the corresponding empirical powers ($\tilde{p}_{12}$) in the case of $S_1^{(1)}(\tau) \neq S_2^{(1)}(\tau)$ under $a^{(1)} = a^{(2)}$ and $\rho^{(1)} = \rho^{(2)}$.

				Clayton			Gumbel			Frank			Marginal	
$\boldsymbol{S}_\tau^{(1)}$	ψ_1^{-1}	ψ_2^{-1}	$\rho^{(k)}$	n_{sim}	n	$\tilde{p}_{12}$	n_{sim}	n	$\tilde{p}_{12}$	n_{sim}	n	$\tilde{p}_{12}$	n_{min}	$\tilde{p}_{12}$
(0.5, 0.1)	1.2	1.2	0.0	2479	2482	80.2	2481	2482	80.0	2479	2482	80.2	2392	78.5
(0.5, 0.1)	1.2	1.3	0.0	2389	2394	80.1	2394	2394	80.0	2389	2394	80.1	2392	80.2
(0.5, 0.1)	1.2	1.5	0.0	2389	2392	80.2	2392	2392	80.0	2389	2392	80.2	2392	80.2
(0.5, 0.1)	1.2	1.2	0.3	2468	2470	80.0	2456	2456	79.9	2449	2456	80.1	2392	78.9
(0.5, 0.1)	1.2	1.3	0.3	2389	2394	80.1	2394	2394	79.9	2389	2394	80.1	2392	80.1
(0.5, 0.1)	1.2	1.5	0.3	2389	2392	80.2	2392	2392	80.0	2389	2392	80.2	2392	80.1
(0.5, 0.1)	1.2	1.2	0.5	2456	2460	80.2	2438	2438	80.1	2436	2436	80.1	2392	79.2
(0.5, 0.1)	1.2	1.3	0.5	2389	2394	80.1	2390	2394	80.0	2389	2394	80.1	2392	80.1
(0.5, 0.1)	1.2	1.5	0.5	2389	2392	80.2	2390	2392	80.0	2389	2392	80.2	2392	80.1
(0.5, 0.1)	1.2	1.2	0.8	2436	2438	80.1	2416	2416	79.9	2412	2414	80.2	2392	79.6
(0.5, 0.1)	1.2	1.3	0.8	2389	2394	80.1	2392	2392	80.0	2389	2392	80.2	2392	80.1
(0.5, 0.1)	1.2	1.5	0.8	2389	2392	80.2	2392	2392	80.0	2389	2392	80.2	2392	80.1
(0.5, 0.1)	1.5	1.5	0.0	545	550	80.7	542	550	80.7	545	550	80.7	532	79.0
(0.5, 0.1)	1.5	1.6	0.0	529	536	80.4	532	536	80.4	529	536	80.4	532	80.1
(0.5, 0.1)	1.5	1.8	0.0	527	532	80.3	528	532	80.4	527	532	80.3	532	80.3
(0.5, 0.1)	1.5	1.5	0.3	542	548	80.5	540	544	80.4	539	544	80.4	532	79.4
(0.5, 0.1)	1.5	1.6	0.3	529	536	80.4	528	534	80.4	529	534	80.2	532	80.2
(0.5, 0.1)	1.5	1.8	0.3	527	532	80.3	528	532	80.4	527	532	80.3	532	80.3
(0.5, 0.1)	1.5	1.5	0.5	539	546	80.5	535	542	80.6	535	540	80.4	532	79.6
(0.5, 0.1)	1.5	1.6	0.5	529	534	80.2	528	534	80.4	527	534	80.3	532	80.2
(0.5, 0.1)	1.5	1.8	0.5	527	532	80.3	528	532	80.3	527	532	80.3	532	80.3
(0.5, 0.1)	1.5	1.5	0.8	537	542	80.4	532	536	80.6	529	536	80.4	532	80.0
(0.5, 0.1)	1.5	1.6	0.8	527	534	80.3	528	532	80.4	527	532	80.3	532	80.3
(0.5, 0.1)	1.5	1.8	0.8	527	532	80.3	528	532	80.5	527	532	80.3	532	80.4
(0.1, 0.5)	1.2	1.2	0.0	2482	2482	80.0	2482	2482	80.2	2482	2482	80.2	2392	78.3
(0.1, 0.5)	1.2	1.3	0.0	1554	1556	80.2	1554	1556	80.1	1554	1556	80.1	1194	64.9
(0.1, 0.5)	1.2	1.5	0.0	1197	1204	80.3	1197	1204	80.2	1197	1204	80.2	1174	79.3
(0.1, 0.5)	1.2	1.2	0.3	2469	2470	80.1	2454	2456	80.2	2456	2456	80.1	2392	78.8
(0.1, 0.5)	1.2	1.3	0.3	1532	1538	80.3	1513	1518	80.1	1513	1518	80.1	1194	66.9
(0.1, 0.5)	1.2	1.5	0.3	1194	1200	80.3	1195	1196	80.1	1190	1196	80.1	1174	79.5
(0.1, 0.5)	1.2	1.2	0.5	2460	2460	80.2	2438	2438	80.0	2434	2436	80.1	2392	79.2
(0.1, 0.5)	1.2	1.3	0.5	1517	1524	80.2	1484	1486	80.1	1480	1484	80.2	1194	68.3
(0.1, 0.5)	1.2	1.5	0.5	1194	1198	80.3	1187	1190	80.2	1186	1188	80.3	1174	79.7
(0.1, 0.5)	1.2	1.2	0.8	2438	2438	80.0	2416	2416	80.0	2412	2414	80.0	2392	79.5
(0.1, 0.5)	1.2	1.3	0.8	1484	1486	80.2	1437	1444	80.2	1428	1436	80.2	1194	70.4
(0.1, 0.5)	1.2	1.5	0.8	1186	1190	80.4	1181	1182	80.2	1176	1182	80.2	1174	79.9
(0.1, 0.5)	1.5	1.5	0.0	542	550	80.6	542	550	80.7	542	550	80.7	532	79.2
(0.1, 0.5)	1.5	1.6	0.0	440	446	80.8	440	446	80.5	440	446	80.5	407	76.1
(0.1, 0.5)	1.5	1.8	0.0	341	348	81.1	341	348	80.7	341	348	80.7	275	67.6
(0.1, 0.5)	1.5	1.5	0.3	544	548	80.5	538	544	80.4	538	544	80.5	532	75.5
(0.1, 0.5)	1.5	1.6	0.3	438	442	80.7	429	438	80.8	433	438	80.7	407	76.2
(0.1, 0.5)	1.5	1.8	0.3	337	344	81.0	333	340	81.1	333	340	81.2	275	72.8
(0.1, 0.5)	1.5	1.5	0.5	540	546	80.5	537	542	80.6	533	540	80.3	532	76.5
(0.1, 0.5)	1.5	1.6	0.5	433	440	80.8	426	432	80.6	425	432	80.7	407	77.0
(0.1, 0.5)	1.5	1.8	0.5	335	340	80.9	327	332	80.8	325	332	80.9	275	73.7
(0.1, 0.5)	1.5	1.5	0.8	536	542	80.5	532	536	80.4	532	536	80.6	532	78.3
(0.1, 0.5)	1.5	1.6	0.8	426	432	80.6	419	424	80.6	417	424	80.8	407	78.7
(0.1, 0.5)	1.5	1.8	0.8	328	334	80.9	319	324	80.8	316	322	80.9	275	75.1

The cases of equal effect sizes As seen in Tables 2 and D.1, n_{min} and n_{ind} are mutually approaching to the sample size n from (3.2) regardless of the copula type and its correlations, according as the effect size ratio δ_2/δ_1 is farther from 1. Conversely, when δ_2/δ_1 is near 1, the sample size n is the farthest from n_{min} and n_{ind}. In such a situation, we can say that it is the most reasonable to use the formula (3.2) considering the correlations $\rho^{(k)}$ between the co-primary endpoints. Table D.2 shows the result of one experiment with the same effect sizes under adopting the settings similar to Table 2 (e.g., equal group size ratio $a^{(1)} = 0.5$, baseline survival rates $S_1^{(1)}(\tau) = S_2^{(1)}(\tau) = 0.1$, 0.5, levels of $\rho^{(k)}$, etc.). This indicates that the sample sizes n from (3.2) provides preferably conservative results similarly to Tables 2 and D.1. We consider the ratio n/n_{ind} under $\rho^{(k)} = 0.8$ to see how much the consideration of correlations $\rho^{(k)}$ make the sample size decrease. In the case of Table D.2, the ratios of n/n_{ind} are between 91.7% and 92.1% if $S_j^{(1)}(\tau) = 0.1$ and 95.9% and 96.4% if $S_j^{(1)}(\tau) = 0.5$.

Table D.2: Total numbers of participants (n) calculated from (3.2) and the corresponding empirical powers ($\tilde{p}_{12}$) under $a^{(1)} = a^{(2)}$, $S_1^{(1)}(\tau) = S_2^{(1)}(\tau)$, $\psi_1 = \psi_2$ and $\rho^{(1)} = \rho^{(2)}$.

		(HR)		Clayton			Gumbel			Frank			Marginal	
$S_j^{(1)}(\tau)$	$\rho^{(k)}$	ψ_1^{-1}	ψ_2^{-1}	n_{sim}	n	$\tilde{p}_{12}$	n_{sim}	n	$\tilde{p}_{12}$	n_{sim}	n	$\tilde{p}_{12}$	n_{min}	$\tilde{p}_{12}$
0.1	0.0	1.2	1.2	1540	1544	80.3	1540	1544	80.3	1540	1544	80.3	1177	69.2
0.1	0.0	1.3	1.3	755	762	80.4	755	762	80.4	755	762	80.4	581	69.7
0.1	0.0	1.5	1.5	328	334	81.0	328	334	81.0	328	334	81.0	256	70.6
0.1	0.0	1.7	1.7	198	204	81.8	198	204	81.8	198	204	81.8	157	71.0
0.1	0.0	2.0	2.0	121	126	82.4	121	126	82.4	121	126	82.4	99	72.4
0.1	0.3	1.2	1.2	1510	1516	80.2	1498	1502	80.1	1494	1498	80.4	1177	67.0
0.1	0.3	1.3	1.3	742	748	80.4	734	742	80.5	732	740	80.6	581	67.0
0.1	0.3	1.5	1.5	322	328	81.0	319	324	80.9	317	324	81.0	256	68.0
0.1	0.3	1.7	1.7	194	200	81.5	193	198	81.6	191	198	81.6	157	69.2
0.1	0.3	2.0	2.0	119	124	82.3	118	124	82.7	117	124	82.8	99	70.7
0.1	0.5	1.2	1.2	1484	1488	80.2	1457	1462	80.2	1452	1452	80.2	1177	69.0
0.1	0.5	1.3	1.3	730	734	80.3	715	722	80.6	713	716	80.3	581	69.0
0.1	0.5	1.5	1.5	317	322	80.9	311	316	80.8	310	314	80.9	256	70.0
0.1	0.5	1.7	1.7	191	196	81.2	187	192	81.2	186	192	81.4	157	71.0
0.1	0.5	2.0	2.0	117	122	82.0	114	120	82.4	114	120	82.7	99	72.3
0.1	0.8	1.2	1.2	1408	1412	80.3	1370	1374	80.3	1332	1340	80.4	1177	73.1
0.1	0.8	1.3	1.3	694	698	80.4	672	678	80.4	657	660	80.3	581	73.1
0.1	0.8	1.5	1.5	302	306	80.7	293	298	81.0	284	290	81.0	256	73.9
0.1	0.8	1.7	1.7	183	186	80.9	176	182	81.5	171	176	81.2	157	74.7
0.1	0.8	2.0	2.0	112	116	81.8	108	114	82.7	105	110	82.2	99	75.8
0.5	0.0	1.2	1.2	3140	3144	80.1	3140	3144	80.1	3140	3144	80.1	2397	64.4
0.5	0.0	1.3	1.3	1564	1570	80.4	1564	1570	80.4	1564	1570	80.4	1199	64.6
0.5	0.0	1.5	1.5	692	700	80.7	692	700	80.7	692	700	80.7	537	65.1
0.5	0.0	2.0	2.0	265	274	81.6	265	274	81.6	265	274	81.6	213	67.2
0.5	0.3	1.2	1.2	3113	3116	80.1	3049	3052	80.0	3060	3062	80.1	2397	66.6
0.5	0.3	1.3	1.3	1548	1556	80.3	1519	1524	80.2	1528	1530	80.4	1199	66.7
0.5	0.3	1.5	1.5	686	694	80.7	674	680	80.7	675	682	80.7	537	67.2
0.5	0.3	2.0	2.0	265	272	81.7	259	266	81.5	260	268	81.7	213	69.0
0.5	0.5	1.2	1.2	3087	3092	80.0	2969	2972	80.0	2978	2978	79.9	2397	68.3
0.5	0.5	1.3	1.3	1534	1544	80.4	1481	1484	80.1	1483	1488	80.2	1199	68.2
0.5	0.5	1.5	1.5	684	688	80.5	654	662	80.6	656	664	80.8	537	68.8
0.5	0.5	2.0	2.0	263	270	81.6	253	260	81.6	253	260	81.4	213	70.5
0.5	0.8	1.2	1.2	3006	3014	80.0	2812	2812	80.0	2758	2760	80.0	2397	71.6
0.5	0.8	1.3	1.3	1504	1506	80.2	1397	1404	80.1	1374	1380	80.3	1199	71.7
0.5	0.8	1.5	1.5	668	672	80.5	622	626	80.4	612	616	80.4	537	72.1
0.5	0.8	2.0	2.0	256	264	81.4	239	246	81.6	236	242	81.2	213	73.6

The cases of unequal group size ratio We investigate how the empirical powers vary when the sampling ratio of each group is not equal. Adopting $\rho^{(k)} = 0.8$ of a high correlation, we consider the patterns of the unequal group size ratios of $a^{(1)} = 0.25$, 0.4, 0.6 and 0.75 and the equal effect sizes of $\psi_j^{-1} = 1.2, 1.3, 1.5, 1.7, 2.0$, $j = 1, 2$ with $S_1^{(1)}(\tau) = S_2^{(1)}(\tau) = 0.1$. Table D.3 is a list of the sample sizes n calculated from (3.2) and the corresponding empirical powers $\tilde{p}_{12}$, where n_{sim} and n_{min} are calculated as before. From Table D.3, we can find a tendency that the empirical powers tend to go below the target power $1-\beta$ as $a^{(1)}$ is farther towards the above from 0.5 but to go above $1-\beta$ as $a^{(1)}$ is lower than 0.5. Comparing the values of n with n_{sim}, their differences are not necessarily ignorable if either effect size (δ_1 and/or δ_2) is smaller. Hence, we should attend to use of the formula (3.2) when the group size ratio $a^{(1)}$ is extremely smaller or larger than 0.5. However, because this problem is not limited to the bivariate case and occurs similarly in univariate case (see Table D.4 of the next paragraph), the major causes are in the univariate critical point rather than the calculation of the bivariate correlation.

Table D.3: Total numbers of participants (n) calculated from (3.2) and the empirical powers ($\tilde{p}_{12}$) in the case of $a^{(1)} \neq a^{(2)}$ under $S_1^{(1)}(\tau) = S_2^{(1)}(\tau) = 0.1$, $\psi_1 = \psi_2$ and $\rho^{(1)} = \rho^{(2)} = 0.8$.

	(HR)		Clayton			Gumbel			Frank			Marginal	
$a^{(1)}$	ψ_1^{-1}	ψ_2^{-1}	n_{sim}	n	$\tilde{p}_{12}$	n_{sim}	n	$\tilde{p}_{12}$	n_{sim}	n	$\tilde{p}_{12}$	n_{min}	$\tilde{p}_{12}$
0.25	1.2	1.2	1848	1904	81.4	1810	1860	81.3	1757	1808	81.4	1580	74.1
0.25	1.3	1.3	905	944	81.9	885	920	82.2	861	896	82.0	780	74.5
0.25	1.5	1.5	387	416	83.1	378	404	83.1	371	396	83.0	344	76.0
0.25	1.7	1.7	231	256	84.2	226	248	84.1	220	240	83.6	208	76.4
0.25	2.0	2.0	140	160	85.4	136	156	85.5	134	152	85.2	128	77.0
0.40	1.2	1.2	1451	1478	80.6	1411	1440	80.7	1381	1403	80.7	1227	73.3
0.40	1.3	1.3	717	733	81.3	697	713	81.1	680	693	81.0	607	73.8
0.40	1.5	1.5	308	323	82.2	299	313	82.1	295	305	81.9	265	74.3
0.40	1.7	1.7	186	198	82.9	181	190	82.5	176	185	82.0	162	75.2
0.40	2.0	2.0	114	123	83.5	111	120	83.7	108	118	84.1	100	75.6
0.60	1.2	1.2	1464	1462	79.7	1424	1424	79.6	1388	1387	79.7	1218	72.6
0.60	1.3	1.3	724	720	79.7	701	702	80.1	684	684	79.8	600	72.1
0.60	1.5	1.5	315	315	79.8	306	307	80.4	298	299	80.1	261	72.8
0.60	1.7	1.7	192	192	80.1	185	187	80.6	181	182	80.5	158	72.6
0.60	2.0	2.0	119	120	80.7	114	117	80.1	111	114	81.2	98	73.2
0.75	1.2	1.2	1894	1854	78.9	1839	1806	79.2	1791	1758	79.1	1548	72.0
0.75	1.3	1.3	940	911	78.7	908	886	78.8	880	862	79.0	760	71.7
0.75	1.5	1.5	411	395	78.2	395	384	78.6	384	374	78.8	329	70.8
0.75	1.7	1.7	250	240	77.9	239	232	78.6	234	227	78.6	200	70.9
0.75	2.0	2.0	154	148	77.6	148	144	78.7	144	140	78.7	124	70.8

The univariate cases In the main text, we calculated the solution n_{min} from PS_{min} using the univariate version of our formula (3.2). Here, we will investigate the performance of the univariate formula, preparing Collett-Freedman's formula (Freedman, 1982; Collett, 2003) as the competitor. We consider the patterns of the group size ratios of $a^{(1)} = 0.25$, 0.4, 0.5, 0.6 and 0.75, the effect sizes of $\psi_1^{-1} = 1.2, 1.3, 1.5, 1.7, 2.0$ and $S_1^{(1)}(\tau) = 0.1$ and 0.5. In this paragraph, let n and n_{CF} be the sample sizes calculated from the univariate version of (3.2) and Collett-Freedman's formula, respectively. Table D.4 is a list of n and n_{CF} and their corresponding empirical powers $\tilde{p}_1$. Similarly to the tendency of Table D.3, the empirical powers corresponding to n tend to go below the target power $1-\beta$ as $a^{(1)}$ is higher than 0.5 but to go above $1-\beta$ as $a^{(1)}$ is lower than 0.5. However, the degree of the tendency is slightly smaller than the bivariate case. On the other hand, the empirical powers corresponding to n_{CF} are farther from $1-\beta$ as $a^{(1)}$ gets farther away from 0.5, and the degree of the tendency is not so better. In any case, the sample sizes calculated from the univariate version of (3.2) have better performance than Collett-Freedman's formula in the empirical powers. Similarly to the bivariate case, probably there is some problem to violation of the normality and room for improvement is left, when the group sizes are extremely unbalanced and/or the sample size n calculated from the formula is relatively small.

Table D.4: Total numbers of participants (n and n_{CF}) calculated from the univariate version of (3.2) and Collett-Freedman's formula and their corresponding empirical powers ($\tilde{p}_1$) in the cases of $S_1^{(1)}(\tau) = 0.1$ and 0.5.

		$S_1^{(1)}(\tau) = 0.5$				$S_1^{(1)}(\tau) = 0.1$			
$a^{(1)}$	ψ_1^{-1}	n	$\tilde{p}_1$	n_{CF}	$\tilde{p}_1$	n	$\tilde{p}_1$	n_{CF}	$\tilde{p}_1$
0.25	1.2	3164	80.6	3020	78.7	1580	81.1	1456	77.9
0.25	1.3	1572	80.7	1472	78.1	780	81.6	696	76.9
0.25	1.5	696	80.9	632	77.5	344	82.5	292	76.3
0.25	1.7	428	81.5	376	76.5	208	83.1	172	76.0
0.25	2.0	268	82.1	228	76.2	128	84.0	100	74.6
0.40	1.2	2485	80.1	2442	79.6	1227	80.4	1190	79.2
0.40	1.3	1237	80.5	1210	79.5	607	80.9	582	79.4
0.40	1.5	550	80.5	532	79.6	265	81.5	250	78.7
0.40	1.7	340	80.9	327	79.8	162	82.5	152	79.9
0.40	2.0	212	81.9	205	79.8	100	82.6	92	79.7
0.50	1.2	2392	80.2	2398	80.2	1174	80.2	1178	80.3
0.50	1.3	1194	80.2	1200	80.4	580	80.5	582	80.5
0.50	1.5	532	80.2	538	80.9	254	80.9	256	81.2
0.50	1.7	328	80.5	334	81.3	154	81.1	158	82.1
0.50	2.0	208	81.0	214	82.1	96	81.8	100	83.1
0.60	1.2	2500	80.0	2553	80.7	1218	79.7	1261	81.2
0.60	1.3	1250	79.9	1290	81.1	600	79.8	631	81.6
0.60	1.5	558	79.9	586	81.9	261	80.0	283	83.1
0.60	1.7	345	80.0	370	83.0	158	80.2	176	84.3
0.60	2.0	220	80.8	240	84.2	98	80.6	113	85.8
0.75	1.2	3212	79.7	3376	81.5	1548	79.1	1685	82.4
0.75	1.3	1608	79.7	1726	82.3	760	78.9	857	83.6
0.75	1.5	720	79.5	802	84.2	329	78.8	394	85.7
0.75	1.7	448	79.5	513	85.1	200	78.7	252	87.3
0.75	2.0	285	80.2	340	86.9	124	78.9	165	89.3

E Computational program

We provide the code for the software R (ver 2.15.2) needed to calculate total numbers of participants (n) calculated from (3.2) (R is a free software package that the user can download from `http://www.r-project.org/`). In advance, please install the R of 32-bit version 2.15.2 and then the additional library package `mvtnorm`. For our computational program, the users can download `samplesize2.zip` file from the web site

`http://www.st.hirosaki-u.ac.jp/~sugimoto/samplesize2.zip`

and extract three dll files (named `theta_rho.dll`, `blogrank_stat.dll` and `libgcc_s_dw2-1.dll`) and one R script file (named `samplesize2.R`) from the `samplesize2.zip` file. The R script file consists of the following code and is defining the function named `samplesize.2sep`. Please write this code on the R console by copy and paste:

```
samplesize.2sep<-function(alpha,power,a1,HR,Sc_e,ta,tf,corr,copula){
##### function to obtain theta corresponding to a given rho #####
thetarho <- function(rho0,cmp,Cp){
clayton=c(0,0.1010,0.2080,0.3277,0.4680,0.6415,0.8696,1.1963,1.7353,2.9366,4.6674)
 gumbel=c(1,0.9033,0.8120,0.7249,0.6407,0.5582,0.4759,0.3918,0.3027,0.2003,0.1340)
frank=c(0,-0.7953,-1.6094,-2.4882,-3.4940,-4.7299,-6.3987,-8.9811,-13.943,-28.613,-57.610)
thetas=rbind(clayton,gumbel,frank)
colnames(thetas)=c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95)
theta=thetas[rownames(thetas)==Cp,which.min(abs(as.numeric(colnames(thetas))-rho0))]
if(min(abs(as.numeric(colnames(thetas))-rho0))==0){return(c(theta,rho0))}
theta0=theta
dyn.load("theta_rho.dll")
theta=.Fortran("theta_rho",rho0,as.integer(cmp),theta=theta0)$theta
return(c(theta,arho))}
##### function to solve integral equation (3.3) #####
KBsolution <- function(alpha,power,rho,gamma,rsd,K){
   if(det(rho) <= 0){print("no positive definite");break}
   library(mvtnorm)
   z_a = qnorm(1-alpha); ndel = 0.001; rn = round(runif(1)*1000); cad=1
   KB = qmvnorm(power, corr=rho, tail="lower.tail")$quantile
   for(j in 1:1000){
     set.seed(rn)
     KB1k = KB*gamma + z_a*(rsd[K]*gamma - rsd[1:(K-1)])
     pow1 = pmvnorm(lower=rep(-Inf,K), upper=c(KB1k,KB), corr=rho)[1]
     G = power - pow1; if(abs(G) < 0.00001 & G <= 0){break}
     F1k = rep(0,K-1)
     for(l in 1:(K-1)){
     vndel = rep(0,K-1); vndel[l] = ndel
     F1k[l] = pmvnorm(lower=c(rep(-Inf,l-1),KB1k[l],rep(-Inf,K-l)),
                 upper=c(KB1k+vndel,KB), corr=rho)[1]/ndel }
     FK = pmvnorm(lower=c(rep(-Inf,K-1),KB),upper=c(KB1k,KB+ndel),corr=rho)[1]/ndel
     dG = -t(F1k)%*%gamma - FK
     KB = KB - G/dG
     if(pow1<0.1){cad=cad+0.5; KB=cad*qnorm(sqrt(power))} }
     return(c(KB))}
#### begin: the computation of sample size formula (3.2)
cmn=ifelse(copula=="clayton",1,ifelse(copula=="gumbel",2,ifelse(copula=="frank",3,0)))
taf=c(ta,tf); K=2
theta<-c(); for(k in 1:K){
theta[k]=thetarho(corr[k],cmn,copula)[1]}
##### begin: computation of bivariate logrank statistics
dyn.load("blogrank_stat.dll")
blogrank_stat<-.Fortran("blogrank_stat",a1,HR,Sc_e,taf,theta,as.integer(cmn),
      delta=numeric(2),Rho=matrix(0,ncol=2,nrow=2),rSD=numeric(2))
##### end: computation of bivariate logrank statistics
delta=blogrank_stat$delta
rSD=blogrank_stat$rSD
Za=qnorm(1-alpha); Zb=qnorm(power)
n_single=(Zb+Za*rSD)^2/(delta^2)
Rho=blogrank_stat$Rho
KB=KBsolution(alpha,power,Rho,delta[1]/delta[2],rSD,K)
n=(KB+rSD[2]*Za)^2/(delta[2]^2)
```

```
out<-list(n,ceiling(ceiling(n*a1)/a1),ceiling(ceiling(n_single*a1)/a1))
names(out) <- c("raw_n", "round_n", "single_n")
return(out)
#### end: the computation of sample size formula (3.2)
}
```

Please put the dll files of `theta_rho.dll`, `blogrank_stat.dll` and `libgcc_s_dw2-1.dll` into your R working folder in advance, because these dll files are used in the above function `samplesize.2sep`, where the former two dll files are composed of computational source codes written originally by FORTRAN and the last dll file is not needed if some similar service is already included in user's PC environment.

The function named `samplesize.2sep` has night arguments, `alpha`, `power`, `HR`, `a1`, `Sc_e`, `ta`, `tf`, `corr` and `copula`. The former eight arguments correspond to the notations α, $1-\beta$, $(\psi_1, \psi_2)'$, $a^{(1)}$, $(S_1^{(1)}(\tau), S_2^{(1)}(\tau))'$, τ_{a}, τ_{f} and $(\rho^{(1)}, \rho^{(2)})'$ in this article, respectively, and the last argument provides one selection of Clayton, Gumbel and Frank copulas.

The following example represents an application where $\alpha = 0.025$, $1-\beta = 0.8$, $a^{(1)} = 0.5$, $(\psi_1, \psi_2) = (1.2^{-1}, 1.3^{-1})$, $(S_1^{(1)}(\tau), S_2^{(1)}(\tau)) = (0.6, 0.3)$, $\tau_{\mathrm{a}} = 2$, $\tau_{\mathrm{f}} = 3$, $(\rho^{(1)}, \rho^{(2)}) = (0.8, 0.8)$ and `copula`="clayton":

```
alpha=0.025          # input type I error for one-sided test
power=0.8            # input target power
a1=0.5               # input ratio of participants assigned to group 1
HR=c(1/1.5,1/1.3)    # input hazard ratios between groups for endpoints 1 and 2
Sc_e=c(0.6,0.3)      # input (ta+tf) years survival of the control
ta=2; tf=3           # input ta=entry time; tf=follow-up time
corr=c(0.8,0.8)      # input correlations between two endpoints of groups 1 and 2
copula="clayton"     # input copula model ("clayton", "gumbel","frank")
> samplesize.2sep(alpha,power,a1,HR,Sc_e,ta,tf,corr,copula)
$raw_n
[1] 945.6165
$round_n
[1] 946
$single_n
[1] 682 810
```

This output is composed of three elements. The 1st element

`samplesize.2sep(alpha,power,a1,HR,Sc_e,ta,tf,corr,copula)$raw_n=945.6165`

is the raw value of the total sample size n calculated from (3.2), and the 2nd element

`samplesize.2sep(alpha,power,a1,HR,Sc_e,ta,tf,corr,copula)$round_n=946`

is a rounding value of the raw n by the rule $\lceil \lceil a^{(1)} n \rceil / a^{(1)} \rceil$, where $\lceil x \rceil$ denotes the smallest integer equal to or more than x. The numerical tables of the total sample sizes in the main text and this supplemental material are created by the rounding values of n. The 3rd element

`samplesize.2sep(alpha,power,a1,HR,Sc_e,ta,tf,corr,copula)$single_n=(682,810)`

represents the total sample size required to test the difference between two groups with the same type I and II errors `alpha` and `1-power` when each of two survival endpoints is regarded as a single endpoint. That is, the maximum of the two sample sizes, `max(samplesize.2sep$single_n)=810`, provides the practical solution $\mathrm{PS}_{\mathrm{min}}$.

Finally, we will provide the confirmation of a part of numerical results of this article using this function `samplesize.2sep`. Please see the 15th line from the bottom of Table 2 in the main text or the 4th line from the bottom of Table D.2 in this supplemental material. The setting parameters are $\alpha = 0.025$, $1-\beta = 0.8$, $a^{(1)} = 0.5$, $(\psi_1, \psi_2) = (1.2^{-1}, 1.2^{-1})$, $(S_1^{(1)}(\tau), S_2^{(1)}(\tau)) = (0.5, 0.5)$, $\tau_\mathrm{a} = 2$, $\tau_\mathrm{f} = 3$, $(\rho^{(1)}, \rho^{(2)}) = (0.8, 0.8)$:

```
> alpha=0.025; power=0.8; a1=0.5; HR=c(1/1.2,1/1.2); Sc_e=c(0.5,0.5)
> ta=2; tf=3; corr=c(0.8,0.8)
> samplesize.2sep(alpha,power,a1,HR,Sc_e,ta,tf,corr,"clayton")$round_n
[1] 3014
> samplesize.2sep(alpha,power,a1,HR,Sc_e,ta,tf,corr,"gumbel")$round_n
[1] 2812
> samplesize.2sep(alpha,power,a1,HR,Sc_e,ta,tf,corr,"frank")$round_n
[1] 2760
```

F Application to discrete survival models of the scope in the article

The main text of the article is discussed focusing on the continuous survival model. We here apply the scope of the present article to discrete survival models. In particular, the two levels case of the discrete models is of interest connected to binary data, even if this is appropriate only in some narrow technical sense.

An extension measure of the hazard ratio defined by $\psi_j(t) = \lambda_j^{(2)}(t)/\lambda_j^{(1)}(t)$ in the continuous models is

$$\psi_j(t) = \frac{d\Lambda_j^{(2)}(t)/(1 - d\Lambda_j^{(2)}(t))}{d\Lambda_j^{(1)}(t)/(1 - d\Lambda_j^{(1)}(t))}, \quad j = 1, 2 \tag{F.1}$$

in the form which comprehends both discrete and continuous times T_{ij}^* (see Cox (1975)). If the j^{th} response variables (T_{ij}^*) are binary (constituted by the values of two levels), $\psi_j(t)$ of (F.1) reduces to the odds ratio.

The correlation between cumulative hazard variates,

$$\rho^{(k)} \quad = \quad \frac{\mathrm{cov}\left[\Lambda_1^{(k)}(T_{i1}^*), \Lambda_2^{(k)}(T_{i2}^*)\right]}{\sqrt{\mathrm{var}[\Lambda_1^{(k)}(T_{i1}^*)]}\sqrt{\mathrm{var}[\Lambda_2^{(k)}(T_{i2}^*)]}}, \quad k = 1, 2$$

can be applied to the discrete survival model (see Prentice and Cai (1992)), where

$$\mathrm{cov}\left[\Lambda_1^{(k)}(T_{i1}^*), \Lambda_2^{(k)}(T_{i2}^*)\right] = \int_0^\infty \int_0^\infty S^{(k)}(t_-, s_-) d\Lambda_1^{(k)}(t) d\Lambda_2^{(k)}(s) - 1,$$

and t_- is a time just prior to t. Similarly to the main text, suppose that the joint survival function is generated by $S^{(k)}(t, s) = \mathcal{C}(S_1^{(k)}(t), S_2^{(k)}(s); \theta^{(k)})$ using some copula model $\mathcal{C}(\cdot)$. Hence, if one of survival variables (T_{i2}^*) is discrete, we have

$$\rho^{(k)} = \left(\int_0^\infty \sum_{l=1}^L \mathcal{C}\left(\mathrm{e}^{-t}, \prod_{j=1}^{l-1}(1 - d\tilde{s}_j); \theta^{(k)}\right) dt d\tilde{s}_l - 1\right) / \sqrt{\mathrm{var}[\Lambda_2^{(k)}(T_{i2}^*)]},$$

where $d\tilde{s}_l = d\Lambda_2^{(k)}(l)$ and the range of possible values of T_{i2}^* is $\{1, 2, \ldots, L\}$. In particular, if $L = 2$, that is, T_{i2}^*'s are binary variables, then we can write

$$\rho^{(k)} = \left(\pi_{\star 1}^{(k)} + \int_0^\infty \mathcal{C}\left(\mathrm{e}^{-t}, \pi_{\star 2}^{(k)}; \theta^{(k)}\right) dt - 1\right) / \sqrt{\pi_{\star 1}^{(k)} \pi_{\star 2}^{(k)}}, \tag{F.2}$$

where $\pi_{\star s}^{(k)} = \Pr(T_{i2}^{*} = s \mid g_i = k)$, $s = 1, 2$.

Table F.1 shows how $\rho^{(k)}$ of (F.2) varies corresponding to $\pi_{\star 2}^{(k)} = 0.1, \ldots, 0.9$ under the copula models in cases where one of bivariate data is continuous and another is binary. The values of $\theta^{(k)}$ adopted for Table F.1 are obtained with $\rho^{(k)} = 0.95$ in Figure 1. Even if $\theta^{(k)}$ provides a high value of $\rho^{(k)}$ in bivariate continuous data, we observe that $\rho^{(k)}$ corresponding to the same $\theta^{(k)}$ falls into a smaller value if either of the variables is transformed to binary data.

Table F.1: The values of $\rho^{(k)}$ under the three copula models in the case of continuous survival and binary data for $\pi_{\star 2}^{(k)} = 0.1, 0.2, \ldots, 0.9$.

					$\pi_{\star 2}^{(k)}$				
Copula ($\theta^{(k)}$)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Clayton (4.667)	0.746	0.772	0.746	0.697	0.635	0.562	0.480	0.384	0.266
Gumbel (0.1340)	0.723	0.769	0.761	0.727	0.678	0.616	0.539	0.443	0.315
Frank (9.554×10^{-26})	0.750	0.798	0.785	0.746	0.691	0.624	0.543	0.445	0.314

The moment calculation of the statistic: the case with continuous survival and binary variables. We consider the case where T_{i1} is continuous survival outcome and T_{i2} is a binary variable. So, assume that T_{i2} is uncensored and has the realized value of 1 or 2. Then, note that $d\Lambda_2^{(k)}(1) = \pi_{\star 1}^{(k)}$, $d\Lambda_2^{(k)}(2) = 1$, $y_2^{(k)}(1) = S_2^{(k)}(0) = 1$, $y_2^{(k)}(2) = S_2^{(k)}(1) = 1 - \pi_{\star 1}^{(k)} = \pi_{\star 2}^{(k)}$ and $H_j^L(1) = a^{(1)}a^{(2)}$. Hence, applying these results to (B.1), (B.2) and (B.4), the asymptotic mean and variances for the logrank statistic for the binary endpoint, which is almost the same as Pearson's chi-square statistic, are

$$\begin{aligned} \mu_2(\tau) &= a^{(1)}a^{(2)}\left(\pi_{\star 1}^{(2)} - \pi_{\star 1}^{(1)}\right), \\ V_{22}(\tau) &= a^{(1)}a^{(2)}\left(a^{(2)}\pi_{\star 1}^{(1)}\pi_{\star 2}^{(1)} + a^{(1)}\pi_{\star 1}^{(2)}\pi_{\star 2}^{(2)}\right) \quad \text{and} \\ V_{22}^{0}(\tau) &= a^{(1)}a^{(2)}\pi_{\star 1}^{\mathrm{p}}\pi_{\star 2}^{\mathrm{p}}, \end{aligned} \tag{F.3}$$

where $\pi_{\star 2}^{\mathrm{p}} = a^{(1)}\pi_{\star 2}^{(1)} + a^{(2)}\pi_{\star 2}^{(2)} = 1 - \pi_{\star 1}^{\mathrm{p}}$. Similar argument yields the asymptotic covariance of

$$V_{12}(\tau) = a^{(1)}a^{(2)} \int_0^{\tau} C(t)\frac{S_1^{(1)}(t)S_1^{(2)}(t)}{S_1^{\mathrm{p}}(t)}\left\{a^{(2)}\frac{dA^{(1)}(t,1)}{S_1^{(1)}(t)} + a^{(1)}\frac{dA^{(2)}(t,1)}{S_1^{(2)}(t)}\right\},$$

which can be approximated again using numerical integration, where $dA^{(k)}(t,s)$, $s = 1, 2$ are

$$dA^{(k)}(t,2) = 0 \qquad \text{and} \qquad dA^{(k)}(t,1) = S^{(k)}(dt,1) - \frac{S^{(k)}(t,1)}{S_1^{(k)}(t)}dS_1^{(k)}(t).$$

Numerical study: continuous and binary co-primary endpoints. Suppose that the 1st endpoint (T_{i1}) is continuous survival outcome and the 2nd (T_{i2}) is a binary variable, and prepare typical cases that sample sizes and correlations are equal in the two groups (i.e., $a^{(1)} = 0.5$, $\rho^{(1)} = \rho^{(2)}$). We set baseline quantities for continuous and binary co-primary endpoints by $S_1^{(1)}(\tau)$ and $\pi_{\star 1}^{(1)}$, and provide the correlation ($\rho^{(k)}$) between continuous and binary endpoints by the latent form (binary data can be generated from continuous variables, such as the result of Table F.1). Similarly to Section 4.1, target power $1 - \beta = 0.8$, the significance level $\alpha = 0.025$, and the censoring distribution $C(t)$ with $\tau_{\mathrm{a}} = 2$ and $\tau_{\mathrm{f}} = 3$ for the survival outcome are used.

Table F.2 is a list of the total sample sizes n calculated from (3.2) and the corresponding empirical powers $\tilde{p}_{12}$ based on Monte-Carlo simulation, which are made selecting some combinations from $S_1^{(1)}(\tau) = 0.1$, 0.4, $\pi_{\star 1}^{(1)} = 0.3$, 0.7, latent $\rho^{(k)} = 0$, 0.5, 0.8, $\psi_1^{-1} = 1.5, 2.0$ and $\psi_2^{-1} = 1.5, 2.0, 2.5$. Then, the empirical powers of $\tilde{p}_{12}$ are not smaller than the desirable power of $1 - \beta$, although the empirical powers become slightly larger than $1 - \beta$ as n is smaller. In this meaning, the formula (3.2) gives an aspect of selectable methods in some settings of continuous and binary co-primary endpoints.

Table F.2: The case with continuous survival and binary co-primary endpoints: the 1st endpoint is survival outcome and the 2nd is binary.

		latent	(HR)	(OR)	Clayton		Gumbel		Frank	
$S_1^{(1)}(\tau)$	$\pi_{\star 1}^{(2)}$	$\rho^{(k)}$	ψ_1^{-1}	ψ_2^{-1}	n	$\tilde{p}_{12}$	n	$\tilde{p}_{12}$	n	$\tilde{p}_{12}$
0.1	0.3	0.0	1.5	1.5	1000	0.800	1000	0.800	1000	0.800
0.1	0.3	0.0	1.5	2.0	418	0.809	418	0.809	418	0.809
0.1	0.3	0.0	1.5	2.5	316	0.808	316	0.808	316	0.808
0.1	0.3	0.0	2.0	2.5	234	0.811	234	0.811	234	0.811
0.1	0.3	0.5	1.5	1.5	1000	0.801	1000	0.810	1000	0.801
0.1	0.3	0.5	1.5	2.0	412	0.807	406	0.807	406	0.808
0.1	0.3	0.5	1.5	2.5	312	0.808	306	0.811	306	0.806
0.1	0.3	0.5	2.0	2.5	232	0.807	232	0.811	232	0.809
0.1	0.3	0.8	1.5	1.5	1000	0.802	1000	0.808	1000	0.800
0.1	0.3	0.8	1.5	2.0	406	0.805	400	0.813	400	0.806
0.1	0.3	0.8	1.5	2.5	306	0.806	300	0.807	300	0.807
0.1	0.3	0.8	2.0	2.5	232	0.809	230	0.807	230	0.808
0.1	0.7	0.0	1.5	1.5	852	0.800	852	0.800	852	0.800
0.1	0.7	0.0	1.5	2.0	352	0.808	352	0.808	352	0.808
0.1	0.7	0.0	1.5	2.5	276	0.804	276	0.804	276	0.804
0.1	0.7	0.0	2.0	2.5	172	0.811	172	0.811	172	0.811
0.1	0.7	0.5	1.5	1.5	852	0.800	850	0.802	850	0.803
0.1	0.7	0.5	1.5	2.0	342	0.807	338	0.806	334	0.803
0.1	0.7	0.5	1.5	2.5	272	0.806	268	0.808	268	0.808
0.1	0.7	0.5	2.0	2.5	170	0.809	168	0.808	168	0.810
0.1	0.7	0.8	1.5	1.5	850	0.802	850	0.802	850	0.800
0.1	0.7	0.8	1.5	2.0	330	0.803	322	0.803	316	0.807
0.1	0.7	0.8	1.5	2.5	264	0.805	262	0.809	258	0.806
0.1	0.7	0.8	2.0	2.5	166	0.806	164	0.800	162	0.801
0.4	0.3	0.0	1.5	1.5	1020	0.802	1020	0.802	1020	0.802
0.4	0.3	0.0	2.0	2.0	382	0.810	382	0.810	382	0.810
0.4	0.3	0.5	1.5	1.5	1016	0.803	1008	0.804	1010	0.803
0.4	0.3	0.5	2.0	2.0	380	0.807	376	0.807	376	0.805
0.4	0.3	0.8	1.5	1.5	1012	0.804	1002	0.802	1002	0.802
0.4	0.3	0.8	2.0	2.0	378	0.806	374	0.805	374	0.806
0.4	0.7	0.0	1.5	1.5	888	0.803	888	0.803	888	0.803
0.4	0.7	0.0	2.0	2.0	304	0.807	304	0.807	304	0.807
0.4	0.7	0.5	1.5	1.5	880	0.803	872	0.802	870	0.804
0.4	0.7	0.5	2.0	2.0	302	0.807	298	0.806	296	0.802
0.4	0.7	0.8	1.5	1.5	870	0.804	864	0.803	862	0.802
0.4	0.7	0.8	2.0	2.0	296	0.802	294	0.805	292	0.802

References

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N.(1993). *Statistical Models based on Counting Processes*. Springer-Verlag.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.

Collett, D. (2003). *Modelling Survival Data in Medical Research*, 2nd edition. Chapman & Hall.

Cox, D. R. (1975). Partial likelihood. *Biometrika* **62**, 269-276.

Fleming, T. R. and Harrington, D. P.(1991). *Counting Processes and Survival Analysis.* John Wiley and Sons.

Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Mathematicae* **19**, 194–226.

Freedman, L. S. (1982). Table of the number of patients required in clinical trials using the logrank test. *Statistics in Medicine* **1**, 121-129.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549–555.

Hougaard, P. (1986). A class of multivariate failure time distribution. *Biometrika* **73**, 671–678.

Jung, S.-H. (2008). Sample size calculation for the weighted ranl statistics with paired survival data. *Statistics in Medicine* **27**, 3350–3365.

Oakes, D. and Manatunga, A. K. (1992). Fisher information for a bivariate extreme value distribution. *Biometrika* **79**, 827–832.

Prentice, R. L. and Cai, J. (1992). Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika* **79**, 495–512.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-07-0, URL http://www.R-project.org.

Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in Copula models for bivariate survival data. *Biometrics* **51**, 1384–1399.

Whitt, W.(2007). Proofs of the martingale FCLT. *Probability Surveys* **4**, 268–302.