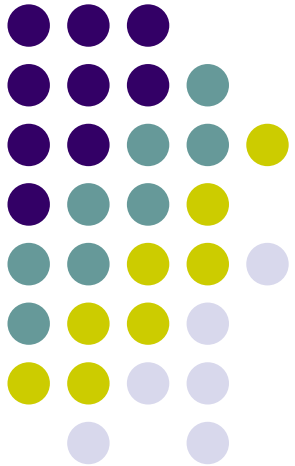
Sample Size Determination for Clinical Trials with Two Correlated Time-to-Event Co-primary Endpoints

The 7th IASC-ARS Joint Taipei Symposium 2011
Academia Sinica, Taipei, Taiwan, December 16-20, 2011

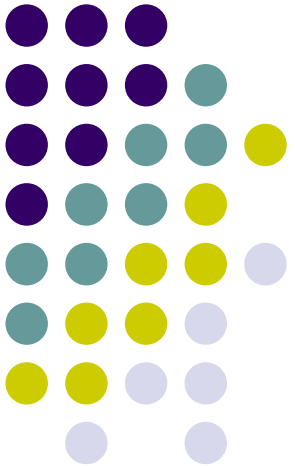
Toshimitsu Hamasaki, PhD
Osaka University Graduate School of Medicine
Scott Evans, PhD
Harvard University School of Public Health
Tomoyuki Sugimoto, PhD
Hirosaki University Graduate School of Mathematical Science
Takashi Sozu, PhD
Kyoto University School of Public Health



This research is financially supported by the following research grants from the MEXT Grant-in-Aid for Scientific Research (C) (No. 23500348), Pfizer Health Research Foundation, Japan and Statistical and Data Management Center of the Adult AIDS Clinical Trials Group grant 1 U01 068634

1. Introduction

Background and Objectives



Clinical Trials with Multiple Endpoints Background

- In clinical trials, historically, ***a single outcome*** is selected as the primary endpoint and is used as the basis for the trial design including sample size determination, as well as for interim monitoring and final analyses.
- Many recent clinical trials become more complex, utilizing more than ***one primary endpoints***
  - **Oncology**
    E1: Time until clinical progression
    E2: Time to death
  - **Prevention of Mother-to-Child HIV/Hepatitis B Transmission**
    E1: Time to infant HIV infection
    E2: Time to Hepatitis B infection
  - **Cardiovascular Disease Therapy**
    E1: Time until the first of MI, Stroke, or death
    E2: Time until hospitalization or death
- The rationale for this is that the assessment of a an intervention using a single endpoint may not provide a comprehensive picture of the intervention's effects.

Strategies for Multiple Endpoints Background

T1) significance on *all* endpoints being sufficient for proof of effect

- Each hypothesis should be rejected at the same significance level
- No adjustment is needed to control type I error
- Type II error increases as the number of outcomes to be tested increases
- "**Multiple Co-Primary Endpoints**" (Hung, Wang, 2009)

T2) significance on *at least one* endpoint being sufficient for proof of effect with a prespecified ordering or non-ordering of outcomes

- Type I error increases as the number of outcomes to be tested increases
- An adjustment to control type I error is required

Hung HMJ, Wang SJ (2009). *J Biopharm Statist* **19**, 1-11.

Arising Natural Questions Background

- How large a sample should be for T1 and T2?
- Is there any considerable overestimation or underestimation in the sample size when the correlation is ignored?
- Is there any considerable reduction or increase in the sample size when the correlation is taken account into the sample size calculation ?

Our Research Focus Objectives

- To discuss the power and sample size determination for superiority comparative clinical trials with **two possibly correlated time-to-events endpoints** to be evaluated as primary variables for the design and analysis, with paying more attention to **T1**
- To consider a **simpler** approach that assumes that the time-to-event outcomes are **exponentially distributed**
- Sugimoto et al (2011) discuss an approach to sizing clinical trials with two correlated time-to-event outcomes based on **the log-rank statistics**.
  - Implementing the method requires technical knowledge, sophisticated programming skill, and expensive computations
- We will focus on hazard ratio : results of difference in hazard rates are very similar as seen in those of hazard ratios

---

Sugimoto T, Hamasaki T, Sozu T (2011). In *Abstract of the 7th International Conference on Multiple Comparison Procedure*, 121, Washington DC, USA, August 29-September 1, 2011.

Co-Primary Endpoints Sample Sizing Related Research

All Continuous Normal Endpoints

Xiong *et al* (2005, *Controlled Clinical Trials*), Sozu *et al* (2006, *Japanese Journal of Biometric Scoiety*), Eaton, Muirhead (2007, *Journal of Statistical Planning and Inference*), Senn, Bretz (2007, Pharmaeutical Statistics), Hung, Wang (2009, *Journal of Biopharmaceutical Statistics*); **Sozu, Sugimoto, Hamasaki (2010, *Statistics in Medicine*; 2011, *Journal of Biopharmaceutical Statistics*)**; **Sugimoto, Sozu, Hamasaki (2011, *Pharmaceutical Statistics*)**; Kordzakhia, Siddiqui, Huque (2010, *Statistics in Medicine*)

All Binary Endpoints

Song (2009, *Computational Statistics and Data Analysis*), **Sozu, Sugimoto, Hamasaki (2010, 2011), Hamasaki, Evans (2011, presented at 2011 Symposium on Applied Statistics)**

All Time-to-Event Endpoints

Sugimoto, Hamasaki, Sozu (2011, presented at MPC2011)

Mixed Endpoints

Sozu, Sugimoto, Hamasaki (2010, presented at IBC2010, mixed continuous and binary endpoints), Sugimoto, Sozu, Hamasaki (2011, presented at MPC2011, mixed binary and time-to-event endpoints)

Outline

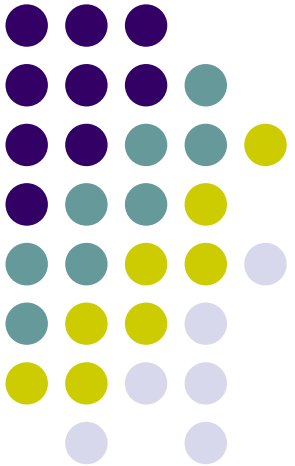
1. Background and Objectives
2. **Comparing log-transformed Hazard ratios (HR) from Two Correlated Exponential Time-to-Event Endpoints**
   - **Statistical Settings**
   - **Conjunctive Power and Sample Size Calculation**
     Without Censoring/Limited Recruitment and Censoring
3. **Behaviors of Sample Size and Empirical Power**
   - **Bivariate Exponential Distributions**
     Clayton Copula/Positive Stable Copula/Fatal-Shock Model
4. **Further Developments**
5. **Summary**

*** Result for difference in hazard rates is available.**

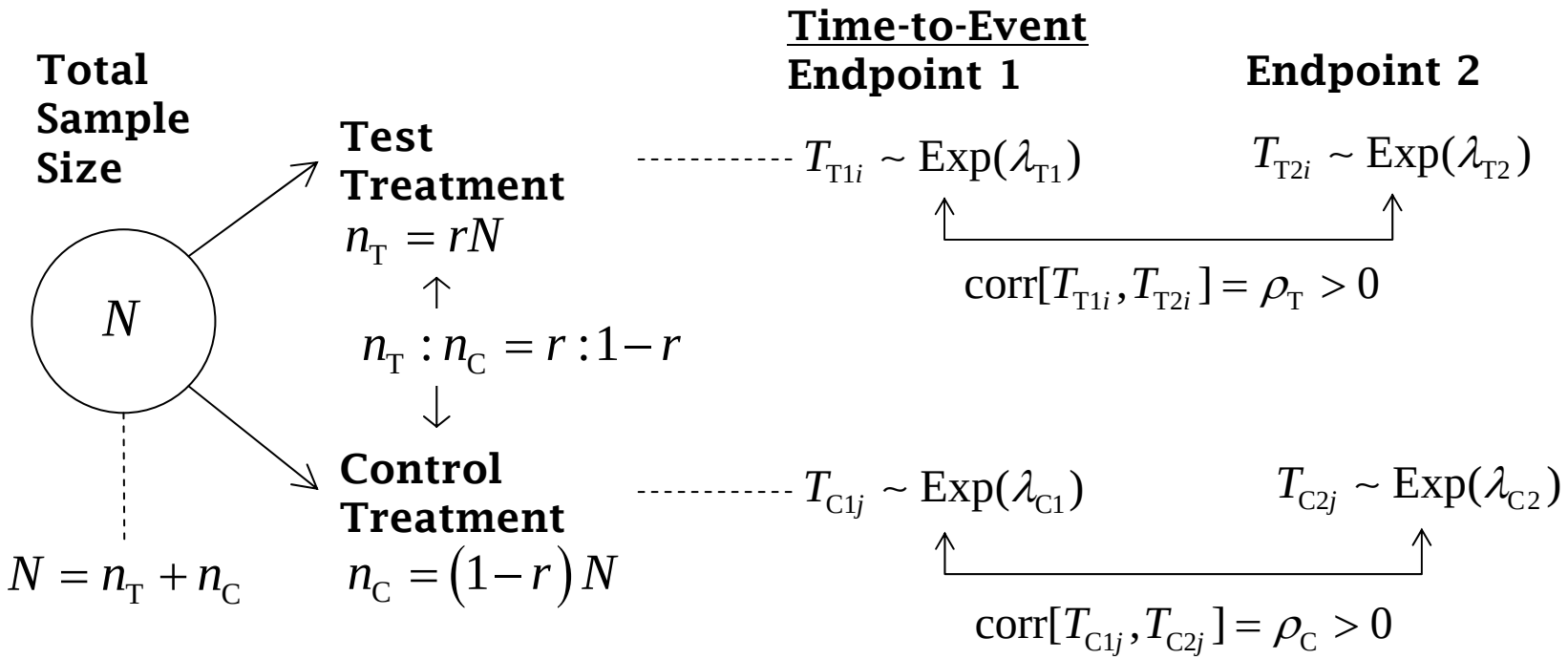
2. Required Sample Size to Compare Hazard Ratio from Two Correlated Exponential Time-to-Event Endpoints

Statistical Setting

Conjunctive Power and Sample Size Calculation



Statistical Settings
Trial Design, Endpoints Distribution



- Randomized, control, superiority clinical trials for two treatment comparison with two time to event endpoints
- T_{Tik}, T_{Cjk} follow the exponential distribution with constant hazard rates $\lambda_{Tk}, \lambda_{Ck} (k = 1,2; i = 1,\ldots,n_T; j = 1,\ldots,n_C)$

Statistical Settings
Distribution of log Hazard Ratio (HR)

Assumption

- Participants are followed until the event of interest
- **No participant is lost to follow-up**

Distributions for large sample

- log-transformed hazard rates → **Approximately normal-distributed**

$$\begin{cases} \log\hat{\lambda}_{\mathrm{T}k} \underset{\text{approx}}{\sim} \mathrm{N}\left(\log\lambda_{\mathrm{T}k}, n_{\mathrm{T}}^{-1}\right) \\ \log\hat{\lambda}_{\mathrm{C}k} \underset{\text{approx}}{\sim} \mathrm{N}\left(\log\lambda_{\mathrm{C}k}, n_{\mathrm{C}}^{-1}\right) \end{cases}$$

- log-transformed hazard ratio→ **Approximately normal-distributed**

$$\begin{aligned} \log\hat{\psi}_k &= \log\hat{\lambda}_{\mathrm{T}k} - \log\hat{\lambda}_{\mathrm{C}k} \\ \log\psi_k &= \log\lambda_{\mathrm{T}k} - \log\lambda_{\mathrm{C}k} \end{aligned} \quad \Rightarrow \quad \begin{aligned} \log\hat{\psi}_1 &\underset{\text{approx}}{\sim} \mathrm{N}\left(\log\psi_1, n_{\mathrm{T}}^{-1} + n_{\mathrm{C}}^{-1}\right) \\ \log\hat{\psi}_2 &\underset{\text{approx}}{\sim} \mathrm{N}\left(\log\psi_2, n_{\mathrm{T}}^{-1} + n_{\mathrm{C}}^{-1}\right) \end{aligned}$$

Collett D (2003). *Modelling Survival Data in Medical Research*. 2nd Edition. Chapman & Hall
Gross AJ, Clark VA.(1975). *Survival Distributions* John Wiley & Sons.

Statistical Setting
Joint Distribution of log HRs

Joint distribution of the two log-transformed HRs for large sample

$$(\log\hat{\psi}_1, \log\hat{\psi}_2) \quad \overset{\text{approx}}{\sim} \quad \mathrm{N}_2(\mu, \Sigma)$$

$$\mu = \begin{pmatrix} \log\psi_1 \\ \log\psi_2 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \text{----} \begin{cases} \sigma_k^2 = \dfrac{1}{N}\left(\dfrac{1}{r} + \dfrac{1}{1-r}\right) & k = k' \\ \sigma_{kk'} = \dfrac{1}{N}\left(\dfrac{\rho_{\mathrm{T}}}{r} + \dfrac{\rho_{\mathrm{C}}}{1-r}\right) & k \neq k' \end{cases}$$

Correlation between the two log-transformed HRs for large sample

$$\rho_{HR} = \mathrm{corr}\left[\log\hat{\psi}_1, \log\hat{\psi}_2\right]$$

$$\approx r\rho_{\mathrm{T}} + (1-r)\rho_{\mathrm{C}} \longrightarrow \rho_{HR} = \rho \longleftrightarrow$$

Common correlation

$$\rho = \rho_{\mathrm{T}} = \rho_{\mathrm{C}}$$

Continuous Endpoints

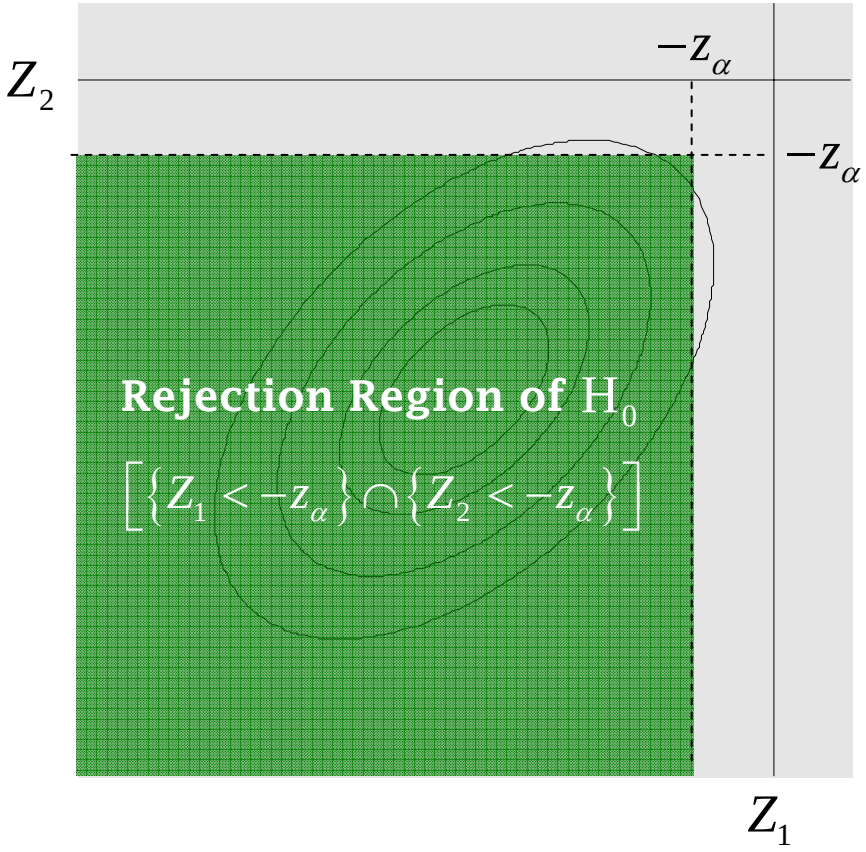
mean difference $\rho_D = \rho$

Binary Endpoints

risk difference $|\rho_{RD}| \leq \rho$

relative risk $|\rho_{RR}| \leq \rho$

Statistical Setting
Hypothesis, Statistics and Rejection Region



Hypothesis for a joint significance

$$\begin{cases} H_1 : \log\psi_1 < 0 & \text{and} \quad \log\psi_2 < 0 \\ H_0 : \log\psi_1 \geq 0 & \text{or} \quad \log\psi_2 \geq 0 \end{cases}$$

Test statistics for hypothesis

$$Z_k = \log\hat{\psi}_k \Big/ \sqrt{\frac{1}{N}\left(\frac{1}{r} + \frac{1}{1-r}\right)}$$

Significant level for hypothesis testing α

z_α is the upper α th percent point of the standard normal distribution

Overall Power and Sample Size Without Censoring

Overall power for showing a joint statistical significance

$$1-\beta = \Pr\left[\bigcap_{k=1}^{2}\left\{Z_k < -z_\alpha\right\}\right]$$

$$\approx \Pr\left[\bigcap_{k=1}^{2}\left\{Z_k^* > c_k\right\}\right]$$

$$Z_k^* = \frac{-\log\hat{\psi}_k + \log\psi_k}{\sqrt{\frac{1}{N}\left(\frac{1}{r}+\frac{1}{1-r}\right)}} \qquad c_k = z_\alpha + \frac{\log\psi_k}{\sqrt{\frac{1}{N}\left(\frac{1}{r}+\frac{1}{1-r}\right)}}$$

"**Conjunctive Power**" or
"**Complete Power**" (Senn, Bretz, 2007)

Sample size

$$N_{NC} = \begin{cases} N & \text{if } N \text{ is an interger} \\ [N]+1 & \text{otherwise} \end{cases}$$

- N is the smallest value satisfying **the overall power**

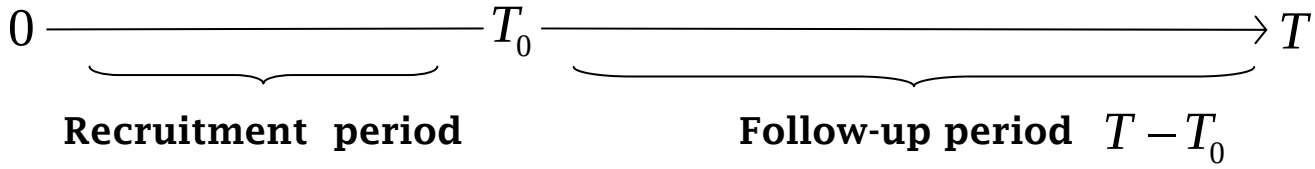
$$1-\beta \longleftarrow \Phi_2\left(-c_1, -c_2 \middle| \rho_{HR}\right)$$

Distribution function of standard bivariate normal distribution

- $[N]$ is the greatest integer less than N

Senn S, Bretz F (2007). *Pharm Statist* **6**, 161-170.

Asymptotic Variance for HR Limited Recruitment and Censoring



- Participants are recruited for study over an interval zero to T_0
- All recruited participants are followed to time of the terminal event or time to $T (T > T_0)$

Asymptotic variance of log-transformed HR for large sample

$$\operatorname{var}\left[\log \hat{\psi}_k\right] \approx \begin{cases} \dfrac{1}{N\phi(\bar{\lambda}_k)}\left(\dfrac{1}{r}+\dfrac{1}{1-r}\right) & \text{Homogeneous variance Null hypothesis} \\ \dfrac{1}{N}\left(\dfrac{1}{r\phi(\lambda_{Tk})}+\dfrac{1}{(1-r)\phi(\lambda_{Ck})}\right) & \text{heterogeneous variance Alterative hypothesis} \end{cases}$$

$$\bar{\lambda}_k = r\lambda_{Tk} + (1-r)\lambda_{Ck} \qquad \phi(\lambda_k) = 1 - \frac{\exp\left(-\lambda_k T + \lambda_k T_0\right) - \exp\left(-\lambda_k T\right)}{\lambda_k T_0}$$

Conjunctive Power and Sample Size Limited Recruitment and Censoring

Over power for showing a joint statistical significance

$$1-\beta = \Phi_2\left(-c_1, -c_2 \middle| \rho_{HR}\right)$$

$$c_k = \frac{\left(z_\alpha \sqrt{\frac{1}{N\phi(\bar{\lambda}_k)}\left(\frac{1}{r}+\frac{1}{1-r}\right)} + \log\psi_k\right)}{\sqrt{\frac{1}{N}\left(\frac{1}{r\phi(\lambda_{Tk})}+\frac{1}{(1-r)\phi(\lambda_{Ck})}\right)}}$$

$$\frac{\lambda_{Tk}^2}{r}+\frac{\lambda_{Ck}^2}{1-r} \geq \bar{\lambda}_k^2\left(\frac{1}{r}+\frac{1}{1-r}\right)$$

$$c_k' = z_\alpha + \frac{\log\psi_k}{\sqrt{\frac{1}{N}\left(\frac{1}{r\phi(\lambda_{Tk})}+\frac{1}{(1-r)\phi(\lambda_{Ck})}\right)}}$$

Sample size

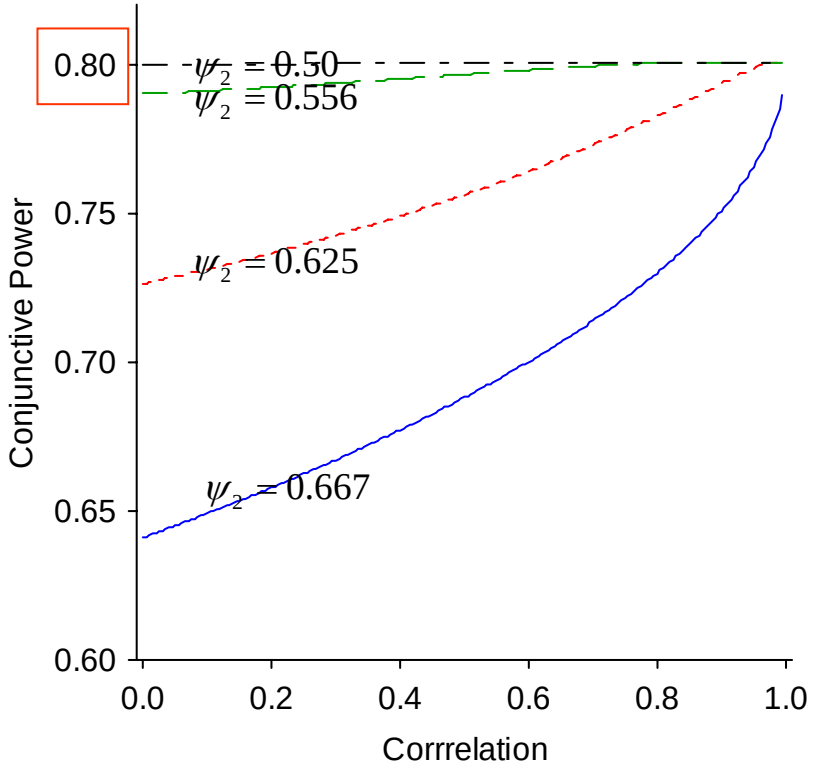
$$N_{CN} = \begin{cases} N & \text{if } N \text{ is an interger} \\ [N]+1 & \text{otherwise} \end{cases}$$

Simplified Sample size

$$N_{CN}^* = \begin{cases} N & \text{if } N \text{ is an interger} \\ [N]+1 & \text{otherwise} \end{cases}$$

→ Improving the approximation

Conjunctive Power
Limited Recruitment and Censoring



$T_0 = 2.0 \quad T = 5.0$

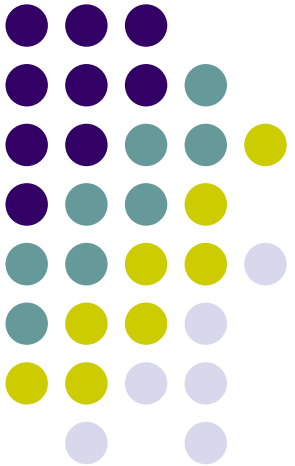
$\psi_1 = 0.667 \quad \lambda_{C1} = 0.5 \quad \lambda_{C2} = 0.5$

$\alpha = 0.025 \quad 1-\beta = 0.8 \quad r = 0.5$

- The overall power increases as the correlation toward one.
- The lowest overall power is when the correlation is zero and the two hazard ratios are equal, with equal hazard rates between control groups

3. Behaviors of Sample Size and Empirical Power

Bivariate Exponential Distributions
Sample Size Behavior
Empirical Power for Log-Rank Test



Models for Correlation
Bivariate Exponential Distributions

1. Clayton Copula Model (Clayton, 1976)

$$S_0(u,v;\theta) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \quad 0 \le \theta$$

θ : Association Parameter

- Times are **positively** associated $0 \le \rho < 1$
- **Late** dependency

2. Positive Stable Copula Model (Hougaard, 1984)

$$S_0(u,v;\theta) = \exp[-\{(-\log u)^{1/\theta} + (-\log v)^{1/\theta}\}^{\theta}] \quad 0 \le \theta \le 1$$

θ : Association Parameter

- Times are **positively** associated $0 \le \rho < 1$
- **Early** dependency

3. Fatal-Shock Model/Marshall-Olkin's Model (Marshall-Olkin, 1967)

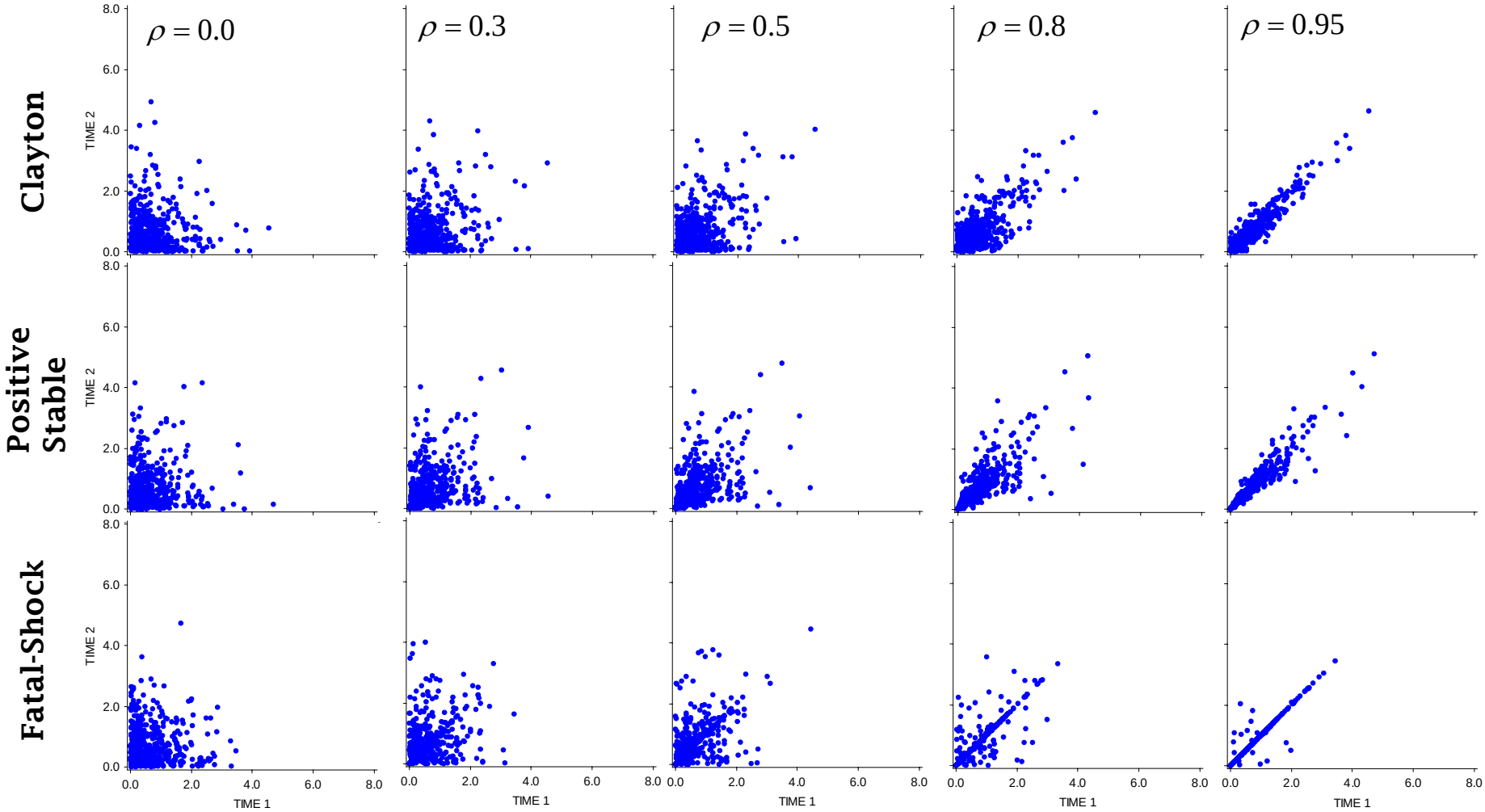
$$S_0(u,v;\lambda_{12}) = \begin{cases} \exp\{-\theta_1 u - (\theta_2 + \theta_{12})v\} & 0 \le u \le v \\ \exp\{-(\theta_1 + \theta_{12})u - \theta_2 v\} & 0 \le v \le u \end{cases}$$

$\theta_1, \theta_2, \theta_{12}$: Hazard Parameter

- The range is **restricted** $0 \le \rho < \min\left(\lambda_1/\lambda_2, \lambda_2/\lambda_1\right)$
- **Linear** dependency

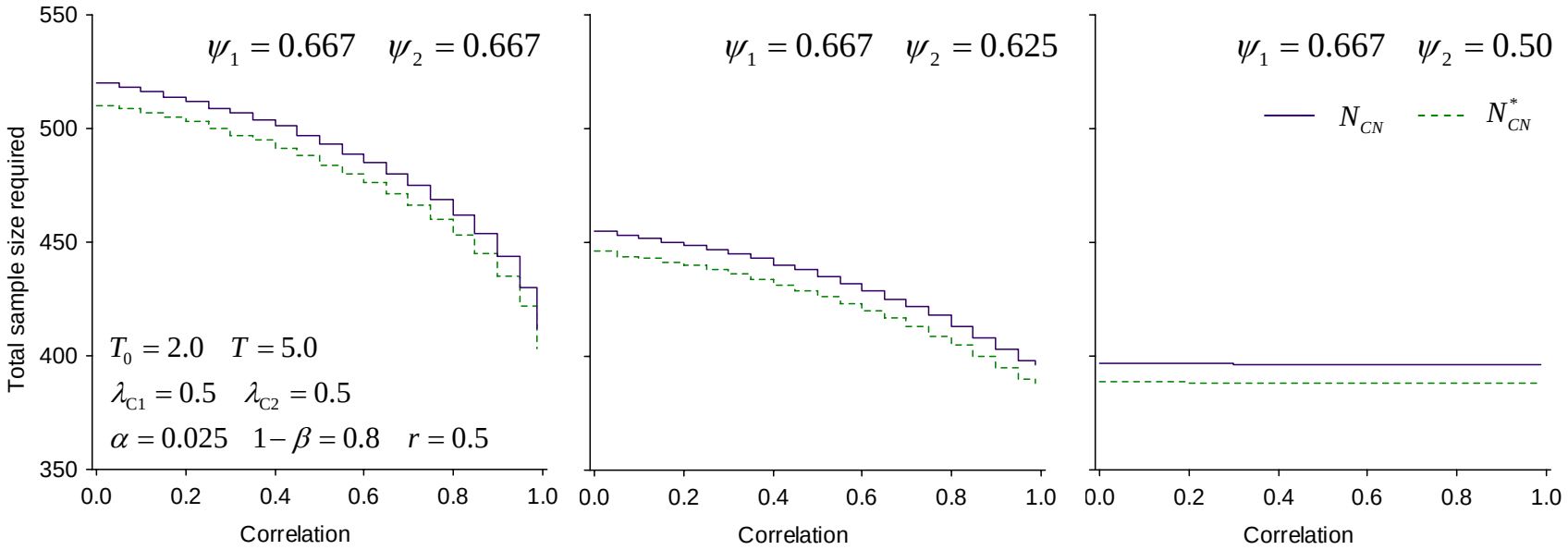
Clayton DG.(1976). *Biometrika* **65**, 141-151. Marshall AW, Olkin I (1967). *J Amer Statist Assoc* **62**, 30-44
Hougaard P.(1984). *Biometrika* **71**, 75-83

Relationship between Two Endpoints Bivariate Exponential Distributions



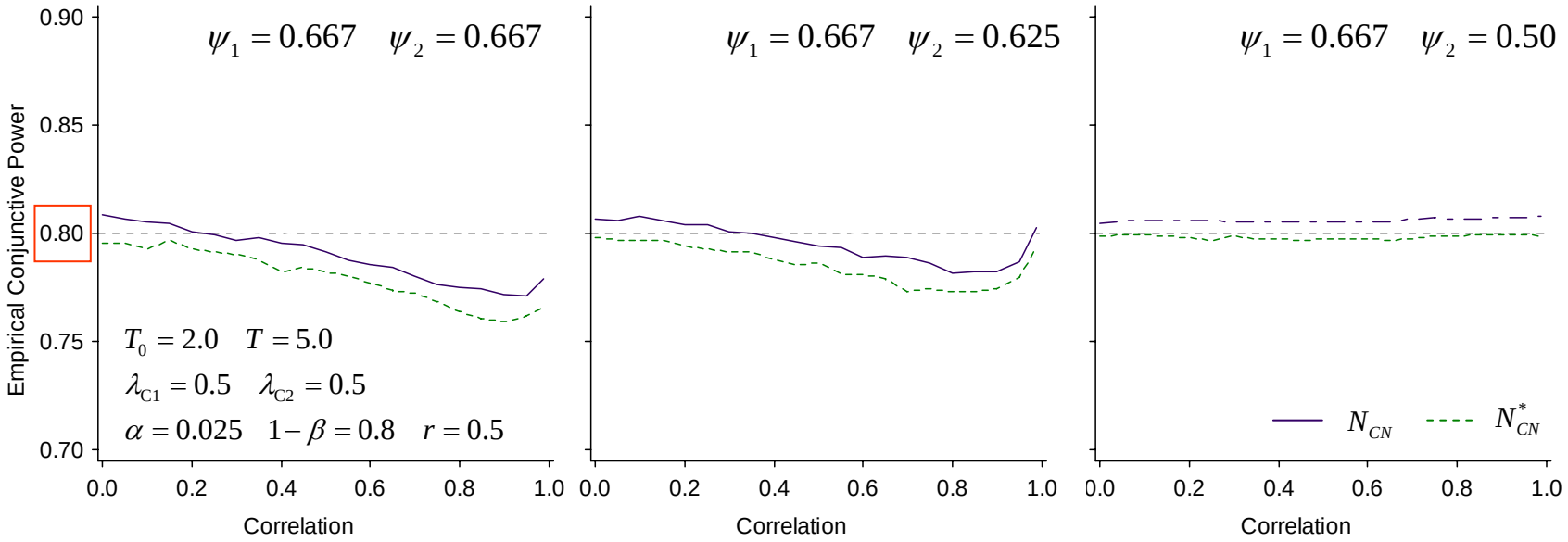
$$\lambda_{T1}/\lambda_{C1} = \lambda_{T2}/\lambda_{C2}$$

Sample Size Behavior
Limited Recruitment and Censoring



- All of the sample sizes decrease as correlation goes toward one. However, the degree of decrease is smaller as the difference between the hazard ratios is larger
- The largest values for all the sample sizes are commonly observed when equal hazard ratio and zero-correlation
- The value of N^*_{CN} is always lager than that of N_{CN}

Empirical Power for Log-Rank Test Clayton Copula Model

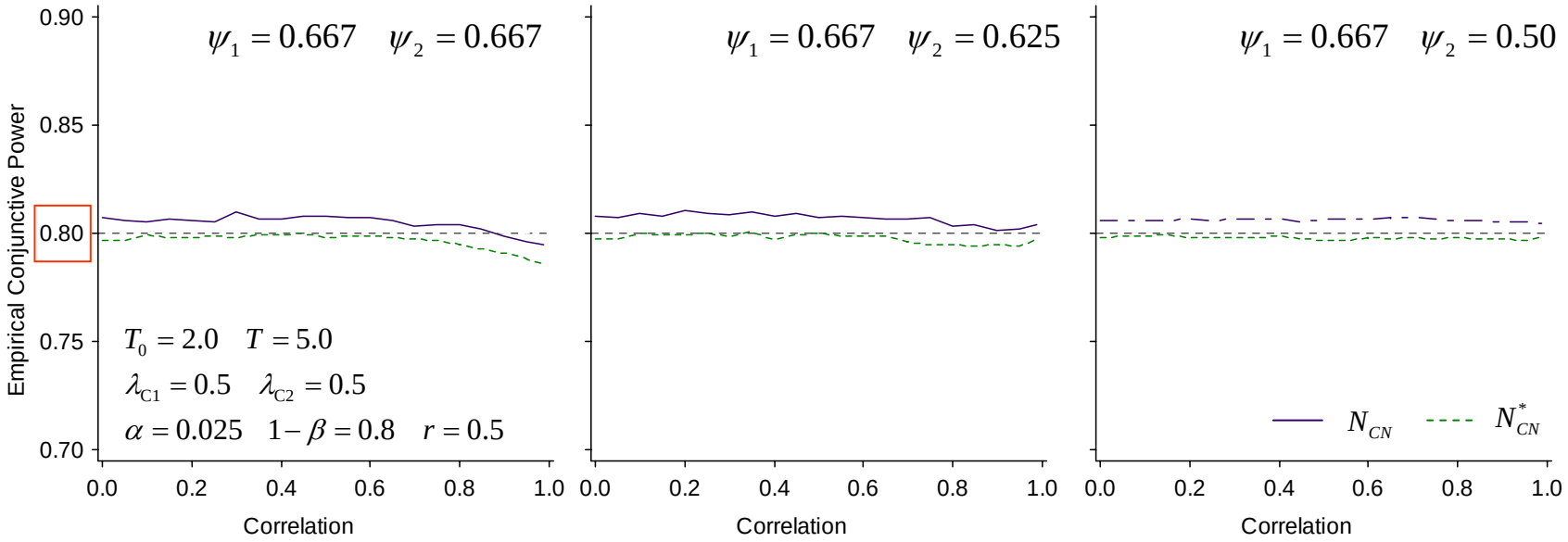


- All of the empirical powers decrease as correlation goes toward one
- In particular the powers are less than the desired power 0.8 as correlation is greater than approximately 0.4 while the empirical powers are greater than the desired power of 0.8 when the correlation is less than around 0.4
- The empirical power of N^*_{CN} is always better than that of N_{CN}

---

* 100,000 Monte-Carlo Trials

Empirical Power for Log-Rank Test Positive Stable Copula Model

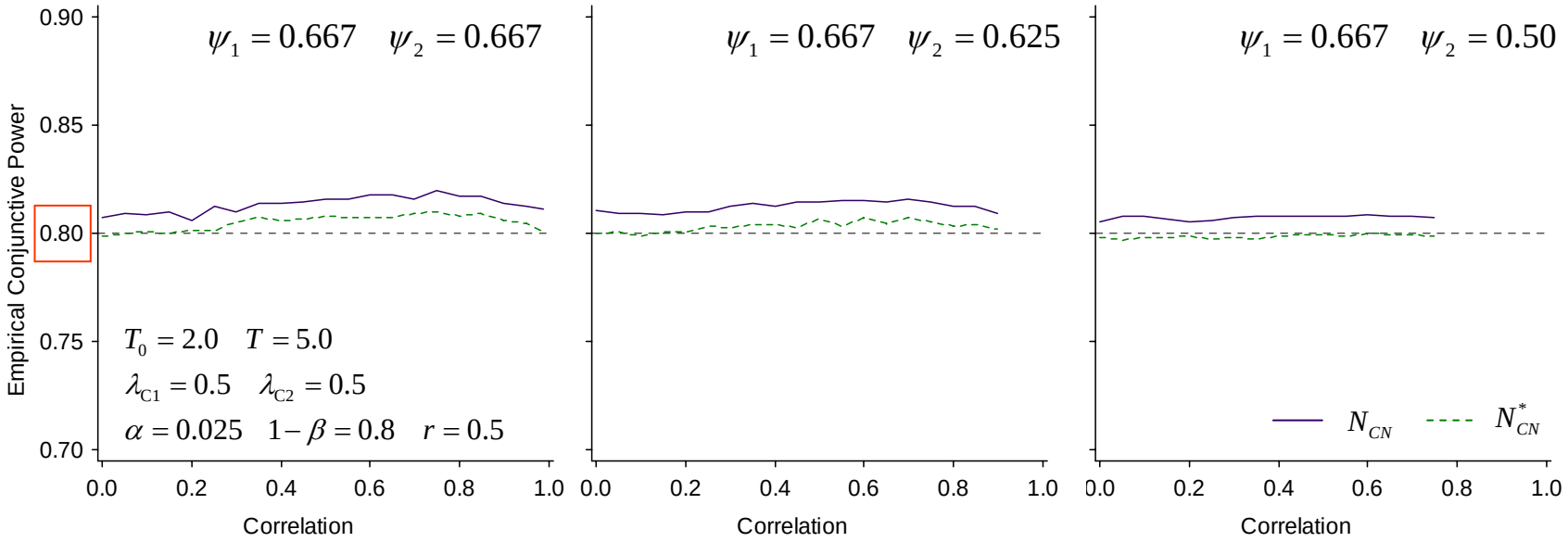


- All of the empirical powers do not much change with correlation and they are attained at the desired power of 0.8
- The empirical power of N^*_{CN} is always slightly larger than that of N_{CN}

---

* 100,000 Monte-Carlo Trials

Empirical Power for Log-Rank Test Fatal-Shock Model



- All of the empirical powers do not much change with correlation and they are attained at the desired power of 0.8
- The empirical power of N^*_{CN} is always slightly larger than that of N_{CN}

---

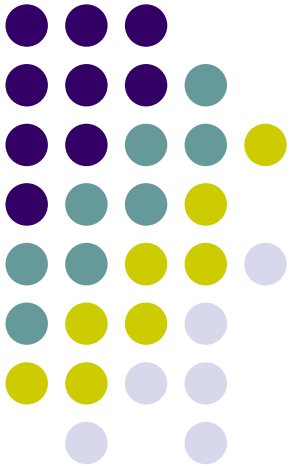
* 100,000 Monte-Carlo Trials

4. Further Developments

At Least One Statistical Significance

Non-Inferiority Hypothesis

Mixed Binary and Time-to-Event Endpoints

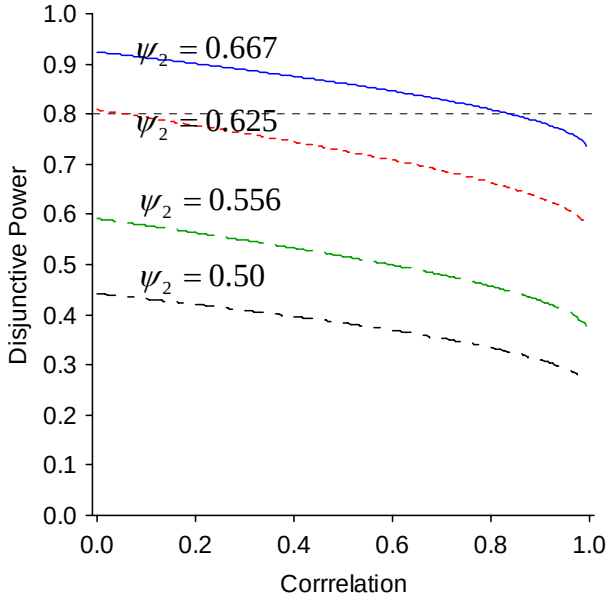


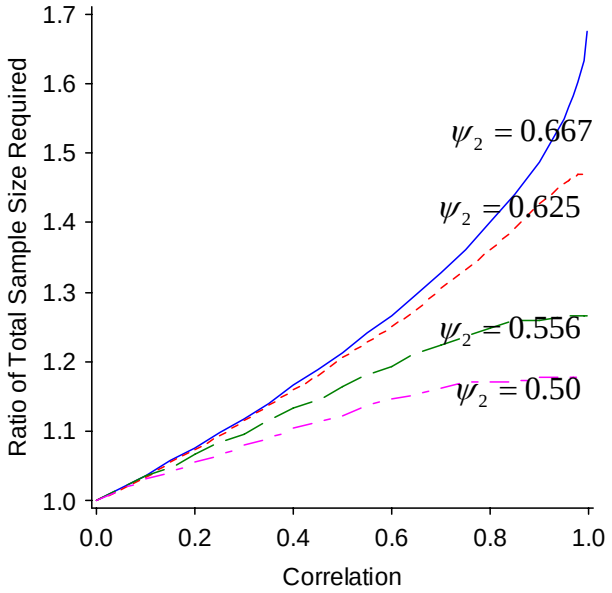
At Least One Statistical Significance Power for Bonferroni Adjustment

Overall power for showing statistical significance for *at least one* endpoint with Bonferroni adjustment

$$1-\beta = 1-\Pr\left[\bigcap_{k=1}^{2}\left\{Z_k > -z_{\alpha/2}\right\}\right]$$

"**Disjunctive power**" or "**Minimal power**" (Senn, Bretz, 2007).





$T_0 = 2.0 \quad T = 5.0$

$\psi_1 = 0.667 \quad \lambda_{C1} = 0.5 \quad \lambda_{C2} = 0.5$

$\alpha = 0.025 \quad 1-\beta = 0.8 \quad r = 0.5$

Non-Inferiority Hypothesis Power and Sample Size

NI hypothesis

$$\begin{cases} \mathrm{H}_1 : \log\psi_1 < \log M_1 & \text{and} \quad \log\psi_2 < \log M_2 \\ \mathrm{H}_0 : \log\psi_1 \geq \log M_1 & \text{or} \quad \log\psi_2 \geq \log M_2 \end{cases} \qquad \begin{cases} M_1 \\ M_2 \end{cases}$$ **Non-inferiority margin**

Test statistics

$$Z_k = \left(\log\hat{\psi}_k - \log M_k\right) \Big/ \sqrt{\frac{1}{N}\left(\frac{1}{r} + \frac{1}{1-r}\right)}$$

Overall power for showing a joint statistical significance (Heterogeneous variance)

$$1-\beta = \Phi_2(-c_1, -c_2 \,|\, \rho_{HR})$$

$$c_k = z_\alpha + \frac{\log\psi_k - \log M_k}{\sqrt{\frac{1}{N}\left(\frac{1}{r\phi(\lambda_{\mathrm{T}k})} + \frac{1}{(1-r)\phi(\lambda_{\mathrm{C}k})}\right)}}$$

Binary and Time-to-Event Outcomes Correlation

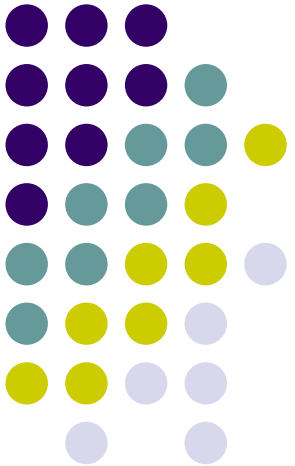
Correlation between hazard ratio and relative risk

$$\operatorname{corr}\left[\log\frac{\hat{\lambda}_T}{\hat{\lambda}_C}, \log\frac{\hat{p}_T}{\hat{p}_C}\right] \approx -\frac{(1-r)\rho_T\lambda_T\sqrt{p_T q_T} + r\rho_C\lambda_C\sqrt{p_C q_C}}{\sqrt{\left\{(1-r)\lambda_T^2 + r\lambda_C^2\right\}\left\{(1-r)p_T q_T + r p_C q_C\right\}}}$$

Binary endpoint	$\longleftrightarrow$	**Time-to-Endpoint**
$Y_{Ti} \sim \text{Bin}(n_T, p_T)$		$S_{Ti} \sim \text{Exp}(\lambda_T)$
$Y_{Cj} \sim \text{Bin}(n_C, p_C)$		$S_{Cj} \sim \text{Exp}(\lambda_C)$
$\text{E}[Y_{Ti}] = p_T \quad \text{var}[Y_{Ti}] = p_T q_T$		$\text{E}[S_{Ti}] = \lambda_T^{-1} \quad \text{var}[S_{Ti}] = \lambda_T^{-2}$
$\text{E}[Y_{Tj}] = p_C \quad \text{var}[Y_{Cj}] = p_C q_C$		$\text{E}[S_{Cj}] = \lambda_C^{-1} \quad \text{var}[S_{Cj}] = \lambda_C^{-2}$

- One of issues is how to define the correlation: a use of correlation form the joint distribution as a limiting distribution of Copulas

5. Summary



Summary

- **We described the power and sample size determination for comparative clinical trials with two correlated time-to-event endpoints to be evaluated as primary variables.**
  - A **simpler** approach that assumes that the time-to-event endpoints are exponentially distributed.
  - Displaying significance on both endpoints for proof of an acceptable efficacy profile
  - The method may work when the dependency structure is early or linear one. While a careful use of the method is recommended when the **late high** dependency is observed.

- **Our research is restricted to "*two* treatment comparison and *two* time-to-event endpoints"**
  - The result from two endpoints gains the insight into more than two endpoints
  - The extension of the result to more than two hazard ratios is not difficult although other issues will arise.

Thank you for your kind attention

If you have any questions, please e-mail to

hamasakt@medstat.med.osaka-u.ac.jp