Dept. Biostatistics
Neurostatistics Working Group
Harvard T.H. Chan School of Public Health
May 25, 2017

USING COMPOSITE ENDPOINTS IN CLINICAL TRIALS

jointly with Steve Lagakos, Ranny Dafni, KyungMann Kim, Moisès Gómez-Mateu, Marta Bofill

Outline of the Talk

1. Composite Endpoints in Randomized Clinical Trials
2. Basic formulae for Sample Size and for Composite Endpoints
3. When cannot be used? Difficulties to anticipate Odds Ratios and Hazard Ratios for Composite Endpoints
4. Asymptotic Relative Efficiency (ARE) as alternative
5. **CompARE**: Interactive Web platform for ARE method and Sample Sizes for CE
6. Concluding Remarks

Primary endpoints in RCT

Variable (outcome) measuring the clinical evidence. Key decision for the study because

- efficacy of new treatment
- power
- sample size computation

are based on the primary endpoint

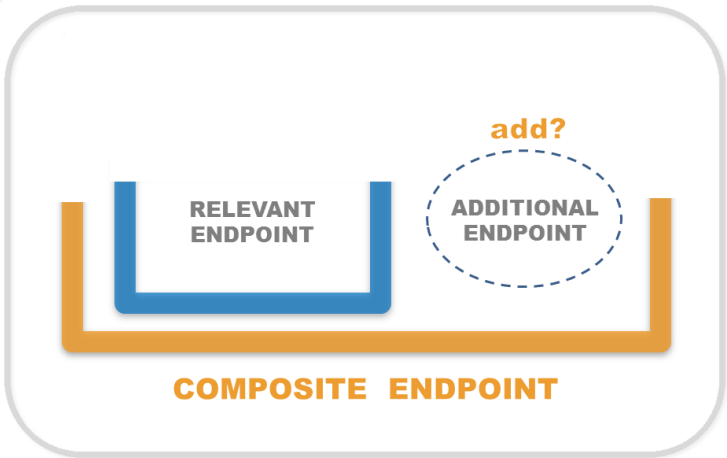
ICH E9 guideline: [1]

If a single primary variable cannot be selected from multiple measurements associated with the primary objective, another useful strategy is to integrate or combine the multiple measurements into a single or composite variable, using a predefined algorithm ... This approach addresses the multiplicity problem without requiring adjustment to type I error.

[1] ICH: Intern. Conf. Harmon. of Tech. Requir. for Registration of Pharmaceuticals for Human Use

COMPOSITE ENDPOINTS: GOOD ALTERNATIVE?

$\mathcal{E}_*$: union of a given set of events $\mathcal{E}_1, ..., \mathcal{E}_k$.



- Achieves a better description of the disease process
- Achieves higher event rates
- Avoids adjustment for multiple comparisons and competing risks
- Inspect Positive CE to determine which components are driving the result
- COULD IMPROVE STATISTICAL EFFICIENCY BY
  - needing smaller sample sizes
  - shorter follow-up times
- HOWEVER POWER IS NOT NECESSARILY REDUCED

MANY AREAS WHERE CE ARE USED

1. **CANCER CLINICAL TRIALS**
   - T_1 time to Disease progression ($\mathcal{E}_1$)
   - T_2 time to Overall survival ($\mathcal{E}_2$)
   - T_* time to PFS: Progression-free survival ($\mathcal{E}_*$)

2. **HIV STUDIES**
   - Y_1 presence/absence of Virological failure ($\mathcal{E}_1$)
   - Y_2 presence/absence of Initiation of new treatment ($\mathcal{E}_2$)
   - Y_* presence/absence of Loss of virological response ($\mathcal{E}_*$)

BINARY CE IN CARDIOVASCULAR STUDIES

Binary Composite Endpoints (BCE)

$Y_* = \mathbf{1}\{\mathcal{E}_* \text{ occurs}\} = \mathbf{1}\{\mathcal{E}_1 \cup \mathcal{E}_2 \text{ occurs}\}$

TAXUS-V[2] TRIAL of Placlitaxel-eluting vs Bare metal stents for coronary artery disease patients.

1. $\mathcal{E}_1 =$ Target-vessel revasc.
2. $\mathcal{E}_2 =$ Death or MI
3. $Y_* =$ Presence of Major adverse cardiac events (MACE)

[2] Stone GW, et al. JAMA. 2005

TIME-TO-EVENT CE IN CARDIOVASCULAR STUDIES

Survival Composite Endpoints (CE)

$T_* = \min\{T_1, T_2, \cdots, T_k\}$ being T_j time from randomization to $\mathcal{E}_j$

EXPEDITION TRIAL[3] of cariporide vs placebo. High risk patients undergoing coronary-artery bypass grafting.

1. $\mathcal{E}_1 =$ Death, T_1 time to Death
2. $\mathcal{E}_2 =$ MI, T_2 time to MI
3. T_* time to earlier event between Death and MI

CONCERN: Positive result ($P = 0.0002$) for CE, driven by a reduction of MI ($P = 0.00005$) but mortality was higher with cariporide

[3] Mentzer et al. Ann Thorac Surg 2008

GOAL: Sample Size for comparing the effect of two interventions, $X = 1$ vs $X = 0$, in superiority clinical trials with CE or BCE

ENDPOINT $\mathcal{E}_1$: T_1 or Y_1

$$\mathcal{H}_1 = \begin{cases} H_{01} : \textbf{NO EFFECT on } \mathcal{E}_1 \\ H_{11} : \textbf{EFFECT on } \mathcal{E}_1 \end{cases}$$

ENDPOINT $\mathcal{E}_2$: T_2 or Y_2

$$\mathcal{H}_2 = \begin{cases} H_{02} : \textbf{NO EFFECT on } \mathcal{E}_2 \\ H_{12} : \textbf{EFFECT on } \mathcal{E}_2 \end{cases}$$

COMPOSITE ENDPOINT $\mathcal{E}_*$: T_* or Y_*

$$\mathcal{H}_* = \begin{cases} H_{0*} : \textbf{NO EFFECT on } \mathcal{E}_* \\ H_{1*} : \textbf{EFFECT on } \mathcal{E}_* \end{cases}$$

Concerns with the interpretation of the effects on $\mathcal{E}_*$

$\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_*$ ARE NOT EQUIVALENT TESTS !!!!!

General sample size for superiority trials

U_{1n} test for $\begin{cases} H_0 : \tau = 0 \equiv \textbf{NO EFFECT on } \mathcal{E}_1 \\ H_1 : \tau < 0 \equiv \textbf{EFFECT on } \mathcal{E}_1 \end{cases}$

Reject H_0 if $U_{1n} \leq C_n$ for some constant C_n (left one-sided alternatives) and assume asymptotic normality (for each n, τ_n is the truth)

$$\frac{\sqrt{n}(U_{1n} - \mu(\tau_n))}{v(\tau_n)} \rightarrow^{\mathcal{L}} N(0,1)$$

Sample Size (n) required to achieve desired power for level α:

$1 - \beta = \pi_n(\tau) \approx 1 - \Phi\left(z_\alpha + \frac{\sqrt{n}(\mu(\tau)-\mu(0))}{v(0)}\right)$

$$\boxed{n \approx \frac{(z_\alpha + z_\beta)^2}{\left(\frac{\mu(\tau)-\mu(0)}{v(0)}\right)^2}}$$

Sample size formula for binary outcomes Y_1

ENDPOINT $\mathcal{E}_1 \equiv Y_1 = \mathbf{1}\{\mathcal{E}_1 \text{ occurs}\}$

$p_1^{(k)} = \text{Prob}\{Y_1 = 1 | X = k\}, k = 0, 1$: probability observing $\mathcal{E}_1$ in group k

$\text{OR}_1 = \frac{p_1^{(1)}/1-p_1^{(1)}}{p_1^{(0)}/1-p_1^{(0)}}$ odds ratio group 1 vs group 0

Assume equal allocation ($n_1 = n_0 = n/2$)

$$\mathcal{H}_1 = \begin{cases} H_{01} : \log(\text{OR}_1) = 0 \\ H_{11} : \log(\text{OR}_1) = \log(o_1) < 0 \end{cases}$$ and U_{1n} is Score Test

$$n_1 = 2\left(\frac{z_\alpha + z_\beta}{\log(o_1)}\right)^2 \cdot \left(\frac{1}{p_1^{(1)} q_1^{(1)}} + \frac{1}{p_1^{(0)} q_1^{(0)}}\right)$$

NEED TO ANTICIPATE VALUES: $(p_1^{(0)}, p_1^{(1)})$ **or** $(p_1^{(0)}, o_1)$

Sample size for time-to-event outcomes T_1

ENDPOINT $\mathcal{E}_1 \equiv$ T_1 time from randomization to $\mathcal{E}_1$

$\lambda_1^{(k)}(t)$: Hazard function for $T_1|X = k$ $(k = 0, 1)$

$\mathrm{HR}_1(\mathrm{t}) = \frac{\lambda_1^{(1)}(t)}{\lambda_1^{(0)}(t)}$. Assume constant $\mathrm{HR}_1(\mathrm{t}) = \mathrm{HR}_1$

Assume equal allocation $(n_1 = n_0 = n/2)$

Assume equal censoring in both groups

$$\mathcal{H}_1 = \begin{cases} H_{01} : \log(\mathrm{HR}_1) = 0 \\ H_{11} : \log(\mathrm{HR}_1(t)) = \log(h_1) < 0 \end{cases}$$

and U_{1n} is Log Rank Test

$$e_1 = 4\left(\frac{z_\alpha + z_\beta}{\log(h_1)}\right)^2 \quad \text{and}^4 \quad n_1 = \frac{2e_1}{p_1^{(0)} + p_1^{(1)}}$$

NEED TO ANTICIPATE VALUES: $(p_1^{(0)}, \mathrm{HR}_1)$

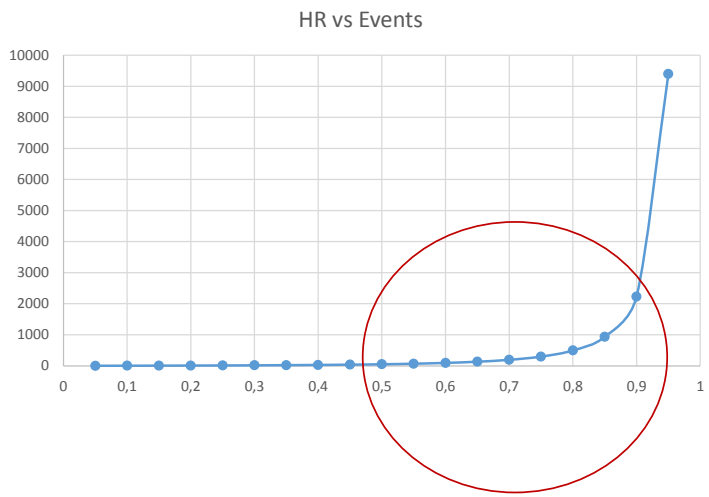
[4] Schoenfeld (Biometrika, 1981)

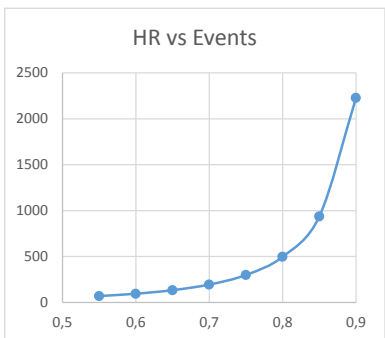
Both formulas have in common $f(x) = \frac{1}{\log(x)^2}$

$$\mathcal{S}(p_1^{(0)}, \mathrm{OR}_1) = 2 \cdot \left(\frac{z_\alpha + z_\beta}{\log(\mathrm{OR}_1)} \right)^2 \cdot \left(\frac{1}{p_1^{(1)} q_1^{(1)}} + \frac{1}{p_1^{(0)} q_1^{(0)}} \right)$$

$$\mathcal{E}(p_1^{(0)}, \mathrm{HR}_1) = 4 \left(\frac{z_\alpha + z_\beta}{\log(\mathrm{HR}_1)} \right)^2$$

HR	Events
0,05	3
0,1	5
0,15	7
0,2	10
0,25	13
0,3	17
0,35	22
0,4	29
0,45	39
0,5	51
0,55	69
0,6	95
0,65	133
0,7	194
0,75	299
0,8	497
0,85	936
0,9	2228
0,95	9400





Small changes in OR_1 and HR_1 (near 1) $\Longrightarrow$ large impact in Sample Size

ANTICIPATED NEEDED VALUES FOR SAMPLE SIZE FOR A CE OR BCE

OUTCOMES: $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_*$

BINARY ENDPOINTS: Y_1, Y_2 and Y_*

SURVIVAL ENDPOINTS: T_1, T_2 and T_*

- Binary and Time-to-event Endpoints
  1. $(p_1^{(0)}, p_2^{(0)})$: Event probabilities in control group
  2. $(\rho^{(0)}, \rho^{(1)})$: Association between $\mathcal{E}_1$ and $\mathcal{E}_2$ for each group
- Binary Endpoints
  3. $(\mathrm{OR}_1, \mathrm{OR}_2)$: Odds Ratios
- Time-to-event Endpoints
  3. $(\mathrm{HR}_1, \mathrm{HR}_2)$: Hazard Ratios
  4. Marginal laws for T_1 and T_2
  5. Joint model for (T_1, T_2) by means of a **Copula** binding the marginal laws for T_1 and T_2

CAN WE USE PREVIOUS FORMULAS TO GET THE SAMPLE SIZE FOR THE CE?

YES, WE CAN, IF WE ANTICIPATE

- Binary Endpoints: $(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho^{(0)}, \rho^{(1)})$
- Time-to-event Endpoints: $(p_1^{(0)}, p_2^{(0)}, \mathrm{HR}_1, \mathrm{HR}_2, \rho^{(0)}, \rho^{(1)})$+ Joint model for (T_1, T_2)+HR_* approximately constant

DIFFICULTIES

1. CORRELATION: Measure of the association between the two events for each group is needed. Not easy to guess!!
2. JOINT LAW FOR (T_1, T_2). In survival models we need as well the joint probabilistic behaviour between T_1 and T_2. Extra distributional assumptions!!
3. HAZARD RATIO: Formulas rely on constant hazard ratios

Composite Binary Endpoint: Treatment effect in terms of Odds Ratio

Endpoint	Probability control group	Probability treat. group	Odds	Odds Ratio
ε_1	$p_1^{(0)}$	$p_1^{(1)}$	$\mathrm{O}_1^{(0)} = p_1^{(0)}/q_1^{(0)}$	OR_1
ε_2	$p_2^{(0)}$	$p_2^{(1)}$	$\mathrm{O}_2^{(0)} = p_2^{(0)}/q_2^{(0)}$	OR_2
ε_*	$p_*^{(0)}$	$p_*^{(1)}$	$\mathrm{O}_*^{(0)} = p_*^{(0)}/q_*^{(0)}$	OR_*

$$\mathrm{OR}_* = \frac{(1+\mathrm{OR}_1\mathrm{O}_1^{(0)})(1+\mathrm{OR}_2\mathrm{O}_2^{(0)}) - 1 - \rho^{(1)}\sqrt{\mathrm{OR}_1\mathrm{OR}_2\mathrm{O}_1^{(0)}\mathrm{O}_2^{(0)}}}{(1+\mathrm{O}_1^{(0)})(1+\mathrm{O}_2^{(0)}) - 1 - \rho^{(0)}\sqrt{\mathrm{O}_1^{(0)}\mathrm{O}_2^{(0)}}} \cdot \frac{1+\rho^{(0)}\sqrt{\mathrm{O}_1^{(0)}\mathrm{O}_2^{(0)}}}{1+\rho^{(1)}\sqrt{\mathrm{OR}_1\mathrm{OR}_2\mathrm{O}_1^{(0)}\mathrm{O}_2^{(0)}}}$$

Odds Ratio's function $\mathcal{OR}(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho^{(0)}, \rho^{(1)})$

Correlation bounds[5]

Correlations $\rho^{(0)}, \rho^{(1)}$ are difficult to guess. Previous studies or pilot studies are a good solution.

Given marginal probabilities $(p_1^{(0)}, p_2^{(0)}, p_1^{(1)}, p_2^{(1)})$ or equivalently given $\theta = (p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2)$ the correlation is bounded and provides a first constrained set of plausible values

$$\boxed{-1 \leq B_L^{(k)}(\theta) \leq \rho^{(k)} = Corr(Y_1^{(k)}, Y_2^{(k)}) \leq B_U^{(k)}(\theta) \leq 1}$$

$$B_L^{(k)}(\theta) = \max\left\{ -\sqrt{\frac{p_1^{(k)} p_2^{(k)}}{(1-p_1^{(k)})(1-p_2^{(k)})}}, -\sqrt{\frac{(1-p_1^{(k)})(1-p_2^{(k)})}{(p_1^{(k)})(p_2^{(k)})}} \right\}$$

$$B_U^{(k)}(\theta) = \min\left\{ \sqrt{\frac{p_1^{(k)}(1-p_2^{(k)})}{(1-p_1^{(k)})(p_2^{(k)})}}, \sqrt{\frac{(1-p_1^{(k)})(p_2^{(k)})}{p_1^{(k)}(1-p_2^{(k)})}} \right\}$$

[5] Sozu *et al.* (Stat Med, 2010)

Correlation bounds when $\rho^{(0)} = \rho^{(1)}$ [6]

Assuming $\rho = \rho^{(0)} = \rho^{(1)}$ simplifies matters but is not always realistic.

Given $\theta = (p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2)$ define

$$B_L(\theta) = \max\{B_L^{(0)}(\theta), B_L^{(1)}(\theta)\} \quad \text{and} \quad B_U(\theta) = \min\{B_U^{(0)}(\theta), B_U^{(1)}(\theta)\},$$

$$\boxed{B_L(\theta) \leq \rho \leq B_U(\theta)}$$

TAXUS-V: Placlitaxel-eluting vs Bare metal stents

1. $Y_1 =$ Occurrence of Target-vessel revasc.
2. $Y_2 =$ Occurrence of Death or MI
3. $Y_* =$ Presence of Major adverse cardiac events (MACE)

$$\theta = (p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2) = (0.173, 0.055, 0.67, 0.81)$$

$$\begin{aligned} B_L(\theta) &= \max\{B_L^{(0)}(\theta), B_L^{(1)}(\theta)\} = \max\{-0.11, -0.08\} = -0.08 \\ B_U(\theta) &= \min\{B_U^{(0)}(\theta), B_U^{(1)}(\theta)\} = \min\{0.53, 0.58\} = 0.53 \end{aligned}$$

$$\boxed{-0.08 \leq \rho \leq 0.53}$$

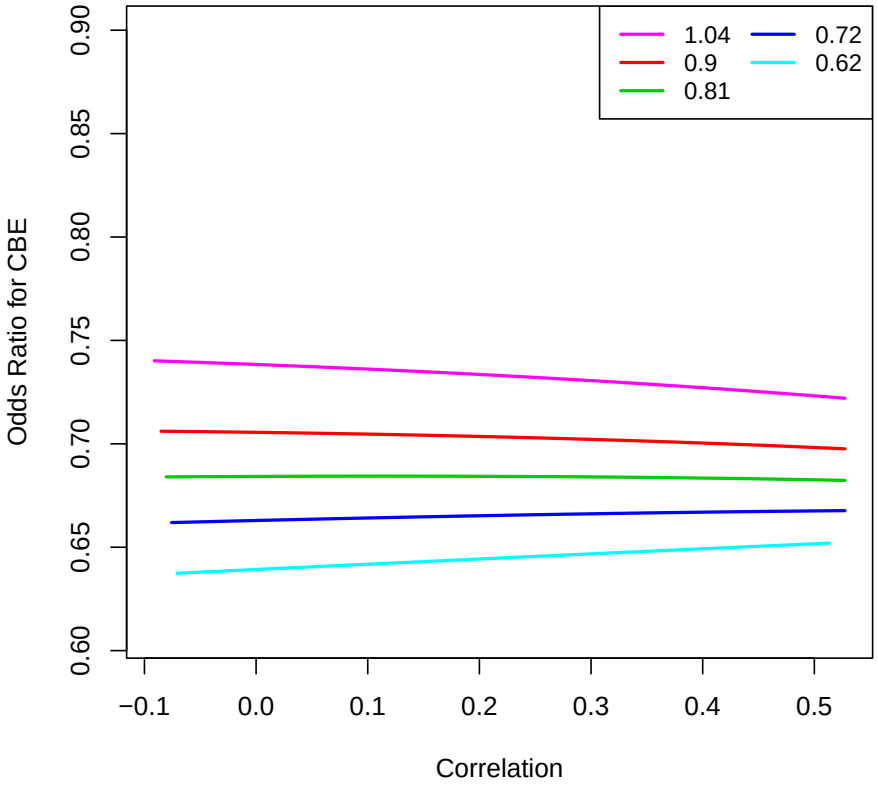
[6] Ongoing work with Marta Bofill

TAXUS-V: $\mathcal{OR}(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho) =$ $\mathcal{OR}(0.173, 0.055, 0.67, \mathrm{OR}_2, \rho)$ VARYING OR_2 AND ρ

$$\min_{B_L(\theta)\le\rho\le B_U(\theta)}\{\mathrm{OR}_*(\rho;\theta)\} \le \mathrm{OR}_*(\theta) \le \max_{B_L(\theta)\le\rho\le B_U(\theta)}\{\mathrm{OR}_*(\rho;\theta)\}$$

$$H_{0*}: \quad \log(\mathrm{OR}_*) = 0$$

$$H_{1*}: \quad \log(\mathrm{OR}_*) < 0$$



OR_2	$\min_\rho\{\mathrm{OR}_*\}$ (ρ)	$\max_\rho\{\mathrm{OR}_*\}$ (ρ)
1.04	0.722 (0.527)	0.740 (−0.091)
0.90	0.698 (0.527)	0.706 (−0.085)
0.81	0.682 (0.527)	0.684 (0.116)
0.72	0.662 (−0.076)	0.667 (0.527)
0.62	0.637 (−0.071)	0.652 (0.513)

Composite Binary Endpoint: Treatment effect in terms of difference of event rates

Endpoint	Probability control group	Probability treat. group	Odds	Odds Ratio
ε_1	$p_1^{(0)}$	$p_1^{(1)}$	$\mathrm{O}_1^{(0)} = p_1^{(0)}/q_1^{(0)}$	OR_1
ε_2	$p_2^{(0)}$	$p_2^{(1)}$	$\mathrm{O}_2^{(0)} = p_2^{(0)}/q_2^{(0)}$	OR_2
ε_*	$p_*^{(0)}$	$p_*^{(1)}$	$\mathrm{O}_*^{(0)} = p_*^{(0)}/q_*^{(0)}$	OR_*

$$
\begin{aligned}
{}^{7}p_*^{(k)} &= 1 - q_1^{(k)} q_2^{(k)} - \rho^{(k)} \sqrt{p_1^{(k)} p_2^{(k)} q_1^{(k)} q_2^{(k)}} \\
\Delta p_* &= p_*^{(1)} - p_*^{(0)}
\end{aligned}
$$

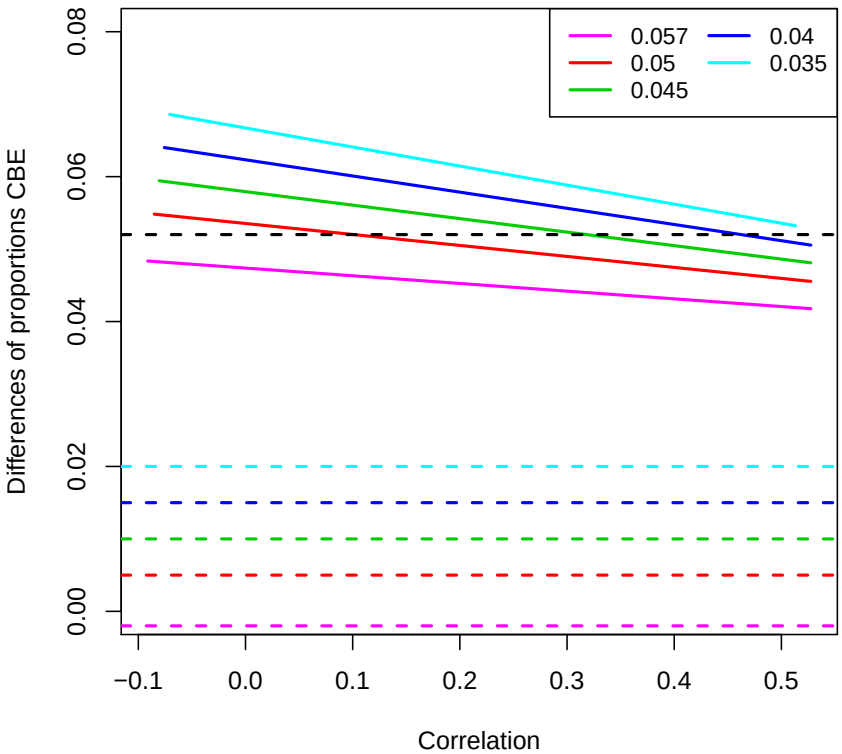
Event Rates Difference $\Delta p_*(p_1^{(0)}, p_2^{(0)}, p_1^{(1)}, p_2^{(1)}, \rho^{(0)}, \rho^{(1)})$

[7] Bahadur (1961)

TAXUS-V: $\Delta p_*(p_1^{(0)}, p_2^{(0)}, p_1^{(1)}, p_2^{(1)}, \rho) =$ $\Delta p_*(0.173, 0.055, 0.122, p_2^{(1)}, \rho)$ VARYING $p_2^{(1)}$ AND ρ

For $p_1^{(1)} - p_1^{(0)} = -0.052$ and

$$H_{0*}: \quad \Delta p_* = p_*^{(1)} - p_*^{(0)} = 0$$
$$H_{1*}: \quad \Delta p_* = p_*^{(1)} - p_*^{(0)} < 0$$



$p_1^{(1)} - p_1^{(0)}$	$= -0.052$	
$p_2^{(1)} - p_2^{(0)}$	$\min_\rho\{\Delta p_*\}$ (ρ)	$\max_\rho\{\Delta p_*\}$ (ρ)
0.002	−0.048 (−0.09)	−0.042 (0.53)
−0.005	−0.055 (−0.08)	−0.045 (0.53)
−0.010	−0.059 (−0.08)	−0.048 (0.53)
−0.015	−0.064 (−0.08)	−0.051 (0.53)
−0.020	−0.069 (−0.07)	−0.053 (0.51)

Sample Size for Composite Binary Endpoint

Sample Size's function $\mathcal{S}(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho^{(0)}, \rho^{(1)})$

$$
\begin{aligned}
H_{0*}: &\quad \log(\mathrm{OR}_*) = 0 \\
H_{1*}: &\quad \log(\mathrm{OR}_*) < 0
\end{aligned}
$$

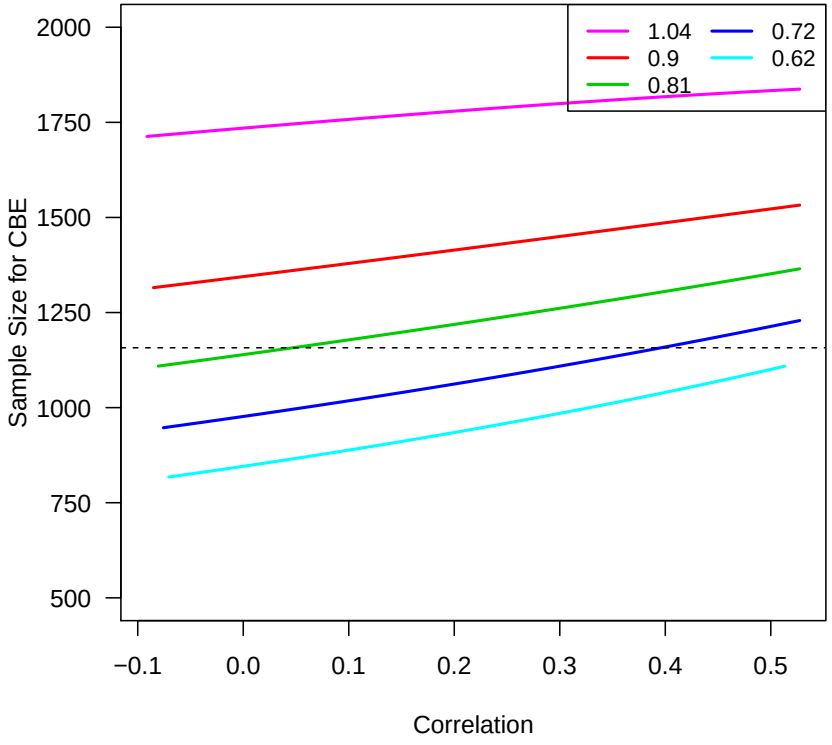
$$
\mathcal{S}(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho^{(0)}, \rho^{(1)}) \quad = \quad 2 \cdot \left(\frac{z_\alpha + z_\beta}{\log(\mathrm{OR}_*)} \right)^2 \cdot \left(\frac{1}{p_*^{(1)} q_*^{(1)}} + \frac{1}{p_*^{(0)} q_*^{(0)}} \right)
$$

For $\rho^{(0)} = \rho^{(1)}$, $\mathcal{S}(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho)$ increases with ρ and is bounded

$$
n_*(B_L(\theta); \theta) \leq n_*(\theta) \leq n_*(B_U(\theta); \theta)
$$

TAXUS-V:

$$\mathcal{S}(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho) = \mathcal{S}(0.173, 0.055, 0.67, \mathrm{OR}_2, \rho)$$



OR_2	$\min_\rho\{n_*\}$ (ρ)	$\max_\rho\{n_*\}$ (ρ)
1.04	1713 (-0.0911)	1838 (0.527)
0.90	1316 (-0.0850)	1533 (0.527)
0.81	1110 (-0.0804)	1365 (0.527)
0.72	948 (-0.0757)	1229 (0.527)
0.62	818 (-0.071)	1109 (0.513)

Conservative approach: Use the upper bound, that is, $n_*(B_U(\theta); \theta)$.

TAXUS-V: Guessing strength of correlation:

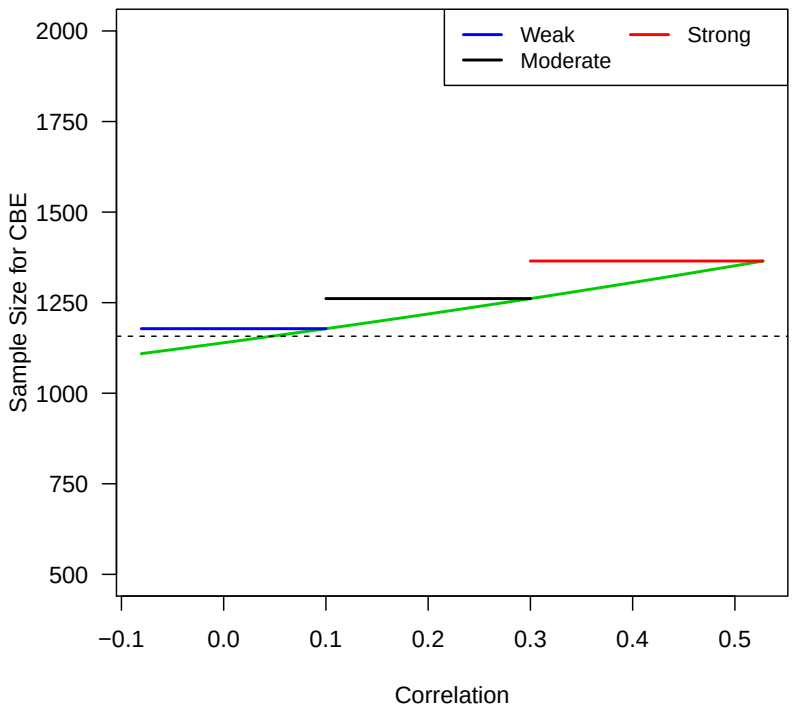
$\mathcal{S}(p_1^{(0)}, p_2^{(0)}, \mathrm{OR}_1, \mathrm{OR}_2, \rho) = \mathcal{S}(0.173, 0.055, 0.67, 0.81, \rho)$

Classify correlation strengths:

- $\rho_W \equiv$ Weak
- $\rho_M \equiv$ Moderate
- $\rho_S \equiv$ Strong

Find bounds for sample size: $n_*^w(B_L(\theta);\theta) \leq n_*(\rho^W;\theta) \leq n_*^w(B_U(\theta);\theta)$

Take the upper bound: $n_*^W(B_U(\theta);\theta)$



- Weak: $-0.081 \leq \rho \leq 0.1$: $n_*^W = 1179$
- Moderate: $0.1 < \rho \leq 0.3$: $n_*^M = 1262$
- Strong: $0.3 < \rho \leq 0.527$: $n_*^S = 1365$

ONGOING

1. Approximation of Odds Ratio, OR_*, or Log Odds ratio for CE in terms of marginal parameters. When is $(OR_1 + OR_2)/2$ a valid summary for OR_*?
2. How to deal with OR_* when $\rho^{(0)} \neq \rho^{(1)}$? How different is the SS if $\rho^{(0)} = 2\rho^{(1)}$?
3. SS under fixed alternatives versus SS under a sequence of contiguous alternatives closer to the null. Theoretical and practical considerations
4. Complete web interface **CompARE** for BCE
   `https://martabofillroig.shinyapps.io/shiny/`

SURVIVAL CE: DEFINING T_* FROM T_1 AND T_2

- T_1 **and** T_2 **via**
  - Marginal densities
  - p_1 and p_2: Probabilities of observing T_1 and T_2 in group 0
  - HR_1 and HR_2 constant relative treatment effects on $\mathcal{E}_1$ and on $\mathcal{E}_2$
- Law of T_*: We need the law of (T_1, T_2).
  - Copula linking the marginal densities
  - ρ: strength of association between T_1 and T_2 (we use Spearman's rank correlation and assume equal for both groups)
- Consider whether T_1 or T_2 include death. Death precludes the observation of the other and is a competing cause. It yields **4 different censoring situations** that have to be worked separately because involve different **marginal or cause-specific hazards**
- $HR_*(t)$ time-dependent even if $HR_1(t) = h_1$ and $HR_2(t) = h_2$

Remark: all the formulas use HR constant, however ...

HR_1 and HR_2 constant $\not\Rightarrow$ $HR_*(t)$ constant.

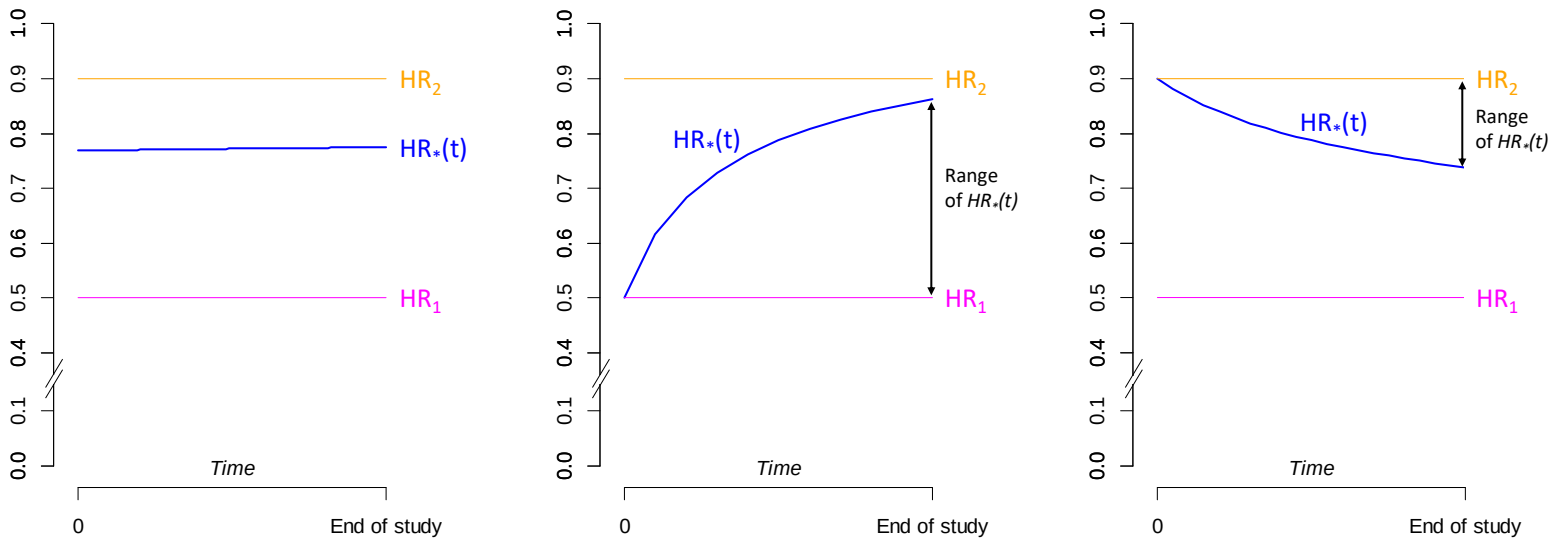


Figure: $HR_*(t)$ for $p_1^{(0)} = 0.05$, $p_2^{(0)} = 0.1$, $HR_1 = 0.5$, $HR_2 = 0.9$.
Left plot: $\rho = 0.1$, exponential for T_1, T_2;
Middle plot: $\rho = 0.5$, exponential for T_1, Weibull increasing hazard rate for T_2;
Right plot: $\rho = 0.5$, Weibull increasing hazard rate for T_1, exponential for T_2.

Sample size for CE T_*

$\lambda_*^{(k)}(t)$: Hazard function for $T_*|X = k$ $(k = 0, 1)$

COMPOSITE ENDPOINT $\mathcal{E}_*$: T_*: $\mathrm{HR}_*(\mathrm{t}) = \frac{\lambda_*^{(1)}(t)}{\lambda_*^{(0)}(t)}$

$$\mathcal{H}_* = \begin{cases} H_{0*} : \mathrm{HR}_*(t) = 1 \\ H_{1*} : \mathrm{HR}_*(t) = h_* < 1 \end{cases}$$ and Log Rank Test

If $HR_*(t) = h_*$ is reasonably constant:

$$e_* = 4\left(\frac{z_\alpha + z_\beta}{\log(h_*)}\right)^2 \quad \text{and} \quad n_* = \frac{2e_*}{p_*^{(0)} + p_*^{(1)}}$$

What to do if $HR_*(t)$ is far from being constant?

- Use alternative measures. **Meaningful option**
- Take advantage of Asymptotic Relative Efficiency (ARE) between using $\mathcal{E}_1$ versus using $\mathcal{E}^* = \mathcal{E}_1 \cup \mathcal{E}_2$

ARE (Asymptotic Relative Efficiency) between using $\mathcal{E}_1$ versus using $\mathcal{E}^* = \mathcal{E}_1 \cup \mathcal{E}_2$ [8]

Research Article

Statistics in Medicine

Received 28 December 2010, Accepted 03 July 2012 Published online in Wiley Online Library

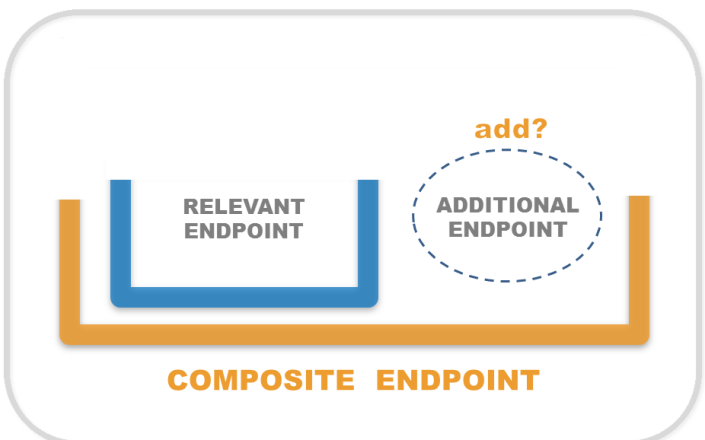
(wileyonlinelibrary.com) DOI: 10.1002/sim.5547

Statistical considerations when using a composite endpoint for comparing treatment groups

Guadalupe Gómez[a*†] **and Stephen W. Lagakos**[b‡]

When comparing two treatment groups in a time-to-event analysis, it is common to use a composite event consisting of two or more distinct outcomes. The goal of this paper is to develop a statistical methodology to derive efficiency guidelines for deciding whether to expand a study primary endpoint from $\mathcal{E}_1$ (for example, non-fatal myocardial infarction and cardiovascular death) to the composite of $\mathcal{E}_1$ and $\mathcal{E}_2$ (for example, non-fatal myocardial infarction, cardiovascular death or revascularisation). We investigate this problem by considering the asymptotic relative efficiency of a log-rank test for comparing treatment groups with respect to a primary relevant endpoint $\mathcal{E}_1$ versus the composite primary endpoint, say $\mathcal{E}_*$, of $\mathcal{E}_1$ and $\mathcal{E}_2$, where $\mathcal{E}_2$ is some additional endpoint. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords: asymptotic relative efficiency; clinical trials; combined outcomes; composite endpoints; log-rank test



[8] Gómez G. and Lagakos S.W. (Stat Med, 2013)

ARE: LogRank for $\mathcal{E}_1$ versus LogRank for $\mathcal{E}^* = \mathcal{E}_1 \cup \mathcal{E}_2$

- $U_{1n} \sim N(0,1)$ under H_0
- $U_{*n} \sim N(0,1)$ under H_0^*
- $U_{1n} \sim N(\mu_1, 1)$ under a sequence of alternatives closer to H_0
- $U_{*n} \sim N(\mu_*, 1)$ under a sequence of alternatives closer to H_0^*

$$\mathrm{ARE} = \left(\frac{\mu_*}{\mu_1}\right)^2$$

$$\mathrm{ARE}\,(U_*, U_1) = \frac{\left(\int_0^1 \log\left\{\frac{\lambda_*^{(1)}(t)}{\lambda_*^{(0)}(t)}\right\} f_*^{(0)}(t)dt\right)^2}{(\log \mathrm{HR}_1)^2 p_1^{(0)} p_*^{(0)}} = \frac{(\mathrm{ALHR}_*)^2 p_*^{(0)}}{(\log \mathrm{HR}_1)^2 p_1^{(0)}}$$

ALHR_*: average log hazard ratio. Could be used as alternative measure

How can we use ARE to get n_* for $\mathcal{E}_*$?[10]

Pitman's Interpretation of ARE

$$\text{ARE} \approx \frac{n_1}{n_*} \Rightarrow n_* \approx \frac{n_1}{ARE}$$

n_1 and n_* required sample sizes for U_{1n} and U_{*n} to have power $1-\beta$ at level α $(0 < \alpha < 1-\beta < 1)$.

Given $(\beta_1, \beta_2, p_1, p_2, \rho)$ and taking into acount if T_1 or T_2 include death, compute $\mathbf{A} = \text{ARE}(\beta_1, \beta_2, p_1, p_2, h_1, h_2, \rho)$.

For given α and power $1-\beta$

(A) If $\text{ARE} \leq 1$, use T_1 with sample size $n_1 = \frac{4(z_\alpha + z_\beta)^2}{(\ln(h_1))^2 p_1^{(0)}}$

(B) If $\text{ARE} > 1$, use T_* with sample size

I $n_* = \frac{4(z_\alpha + z_\beta)^2}{(\ln(h_*))^2 (p_*^{(0)}(t))}$ if $\text{HR}_*(t) \approx h_*$ for all t,

II $n_* = \frac{4(z_\alpha + z_\beta)^2}{\mathbf{A}(\ln(h_1))^2 p_1^{(0)}}$ if $\text{HR}_*(t) \leq h_*$ for all t, not constant

[10] Gómez G. and Gómez-Mateu M. (Sort, 2014)

Web for Sample Sizes for Time-to-Event CE

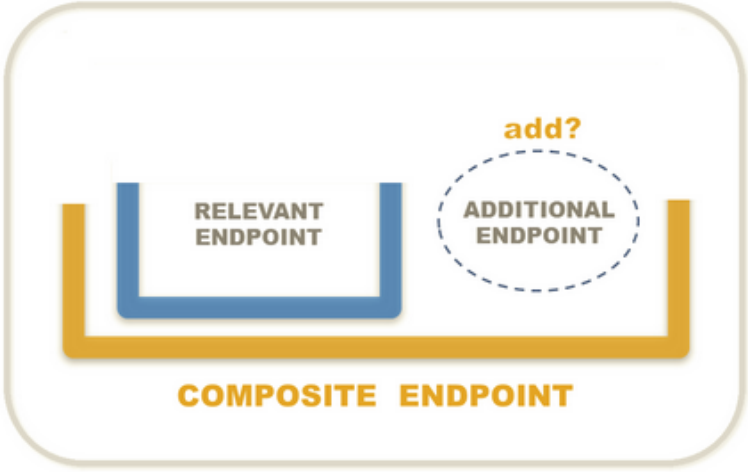
http://cinna.upc.edu:3838/compare/compare_check_2/

Home | ARE value | ARE by correlation (plot) | ARE by correlation (table) | Sample Size | Study of HR*(t) | Help

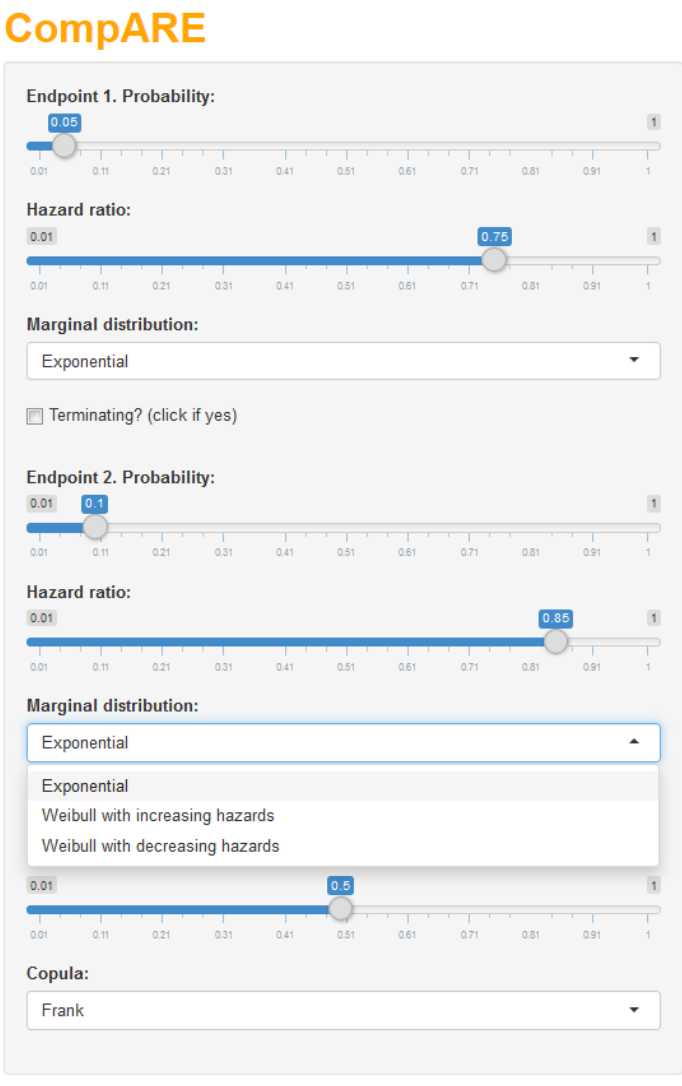
Welcome to Compare platform.

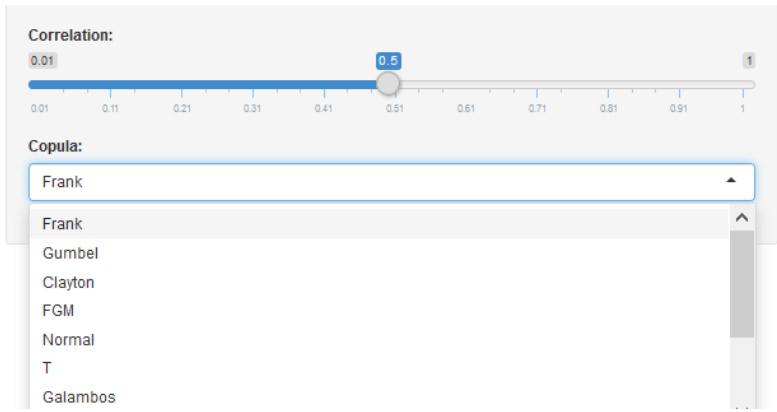
This website will help you to:

* Analyze whether you should use a Composite endpoint as Primary endpoint.
* Sample size for time-to-event data.
* Compare different scenarios depending on your candidate endpoints.
* Get helpful numerical and intuitive graphical results.



Anticipated values





ARE as a function of correlation

CompARE

Endpoint 1. Probability:

0.05 (1)

Hazard ratio:

0.01 — 0.75 — 1

Marginal distribution:

Exponential

☐ Terminating? (click if yes)

Endpoint 2. Probability:

0.01 — 0.1 — 1

Hazard ratio:

0.01 — 0.85 — 1

Marginal distribution:

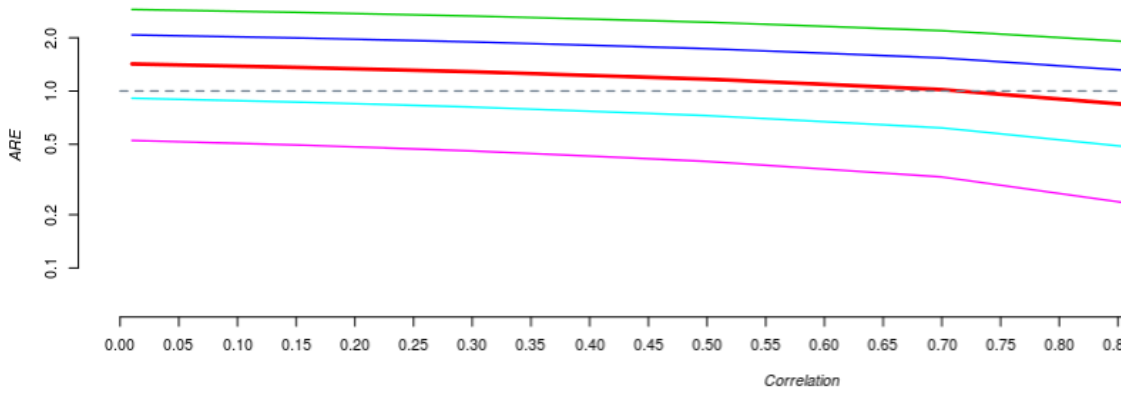
Exponential

☐ Terminating? (click if yes)

Correlation:

0.01 — 0.5 — 1

Home | ARE value | ARE by correlation (plot) | ARE by correlation (table) | Sample Size | Study of HR*(t) | Help



Copula:

Frank

- Frank
- Gumbel
- Clayton
- FGM
- Normal
- T
- Galambos

Sample Size as a function of correlation

Home | ARE value | ARE by correlation (plot) | ARE by correlation (table) | Sample Size | Study of HR*(t) | Help

Total sample size using Endpoint 1:

6811

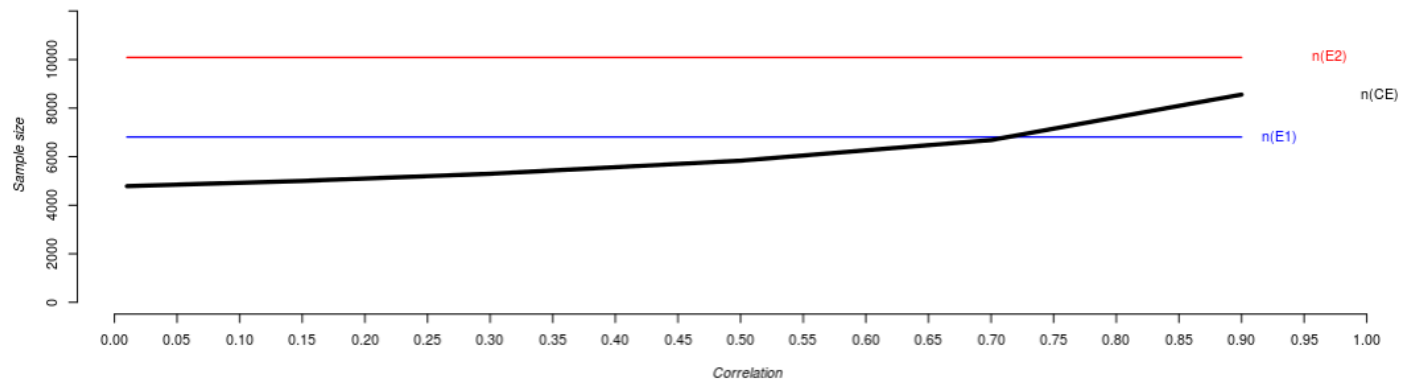
Total sample size using Endpoint 2:

10086

Total sample size using Composite Endpoint:

5834

Sample size depending on different correlations



$\alpha = 0.05,\ 1-\beta = 0.8,\ p_1^{(0)} = 0.05,\ \mathrm{HR}_1 = 0.7 \Rightarrow \mathbf{n_1} = \mathbf{4560}$ patients

If $p_2^{(0)} = 0.05,\ \mathrm{HR}_2 = 0.8 \Rightarrow \mathbf{n}_* = \begin{cases} \mathbf{3699}, & \text{if } \rho = 0.1 \\ \mathbf{4149}, & \text{if } \rho = 0.5 \end{cases}$

Non Proportional Hazards (NPH) are being detected more frequently: Why is so?

- Phase III trials are much larger $\Rightarrow$ more power to detect NPH
- Rare events and small effects with new, better therapies $\Rightarrow$ Composite endpoints used more often $\Rightarrow$ NPH

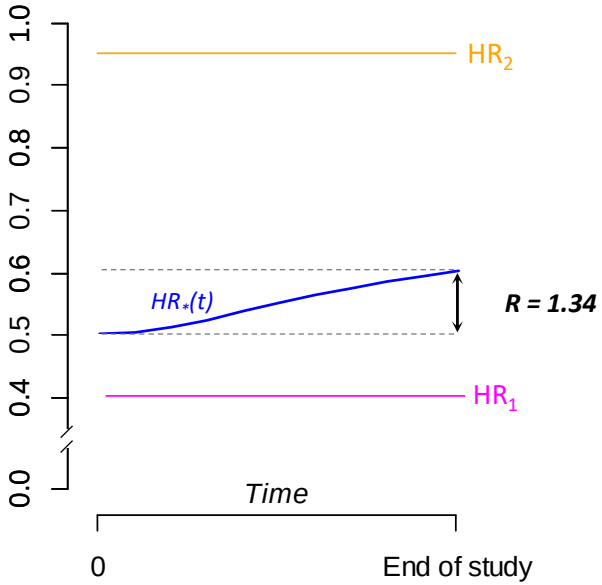
If PH holds, HR would partially capture the relative difference between two survival curves and can be used as a measure to quantify the between-group difference

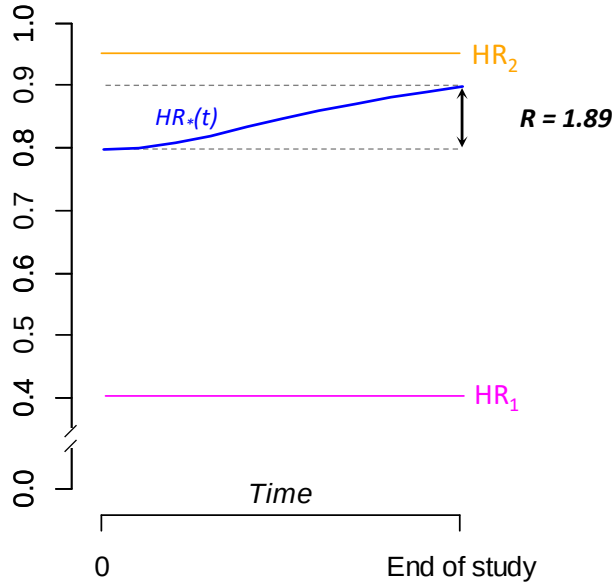
Consequences of NPH:

- If PH is violated, $\mathrm{HR}(t)$ changes over time, the parameter being estimated is not a meaningful measure of the between-group difference, is not the average of the true hazard ratio over time.
- $\mathrm{HR}(t)$ lacks the context to allow to translate the HR into a more understandable clinical benefit.

Departure from constancy of $HR_*(t)$[11]

$R = n_{MHR_*}/n_{aHR_*}$ measures impact on sample size for deviance from being constant.





	Treatment effect				
Laws of each component	$HR_1 = HR_2$	$\lvert HR_1 - HR_2 \rvert = 0.1$	$\lvert HR_1 - HR_2 \rvert = 0.2$	$\lvert HR_1 - HR_2 \rvert = 0.3$	
Both decreasing hazards	0%	0%	0%	0%	**0%**
Both exponential	0%	0%	0%	4%	**0%**
Both increasing hazards	0%	0%	3%	11%	**2%**
Different behaviour on hazards	2%	65%	93%	98%	**60%**
	2%	**43%**	**62%**	**67%**	**41%**

[11] Ongoing work with Moisès Gómez-Mateu and KyungMann Kim

WRAPPING UP

- Composite Endpoints (CE) are very often used as PE in phase 3 RCT
- Phase 3 RCT are usually powered to achieve clinically relevant outcomes and Determination of sample size (SS) is fundamental. It is not an easy task when the PE is a CE
- $HR_*(t)$ is often not constant
- Correlations difficult to guess
- OR_* difficult to get
- ARE as a tool to compute required SS for CE

ONGOING RESEARCH

1. Formulae when $\rho^{(0)} \neq \rho^{(1)}$
2. Average log hazard ratio as alternative summary measures when PH fails
3. Finish web interface **CompARE** for time-to-event CE
4. Unify web interfaces **CompARE** for Binary and Time-to-event CE

THANKS TO MY COAUTHORS AND TO STEVE


THANK
YOU

See you in Barcelona in JULY 2018 !!!!!!