PLANNING CLINICAL TRIALS WITH COMPOSITE ENDPOINTS

4th February, 2016

Outline of the Talk

1. Overview Clinical Trials
2. Overview Composite endpoints
3. Asymptotic relative Efficiency (ARE) to guide the choice of the primary endpoint
4. **CompARE**: Web platform to facilitate the decision between CE and RE as the primary endpoint of the RCT

Overview Clinical Trials

Clinical trial: experiment in clinical research on human participants and designed to answer specific questions about biomedical or behavioral interventions.

It aims to ensure the scientific validity and reproducibility of the results by generating high quality scientific evidence.

- It can be considered the experimental step of the scientific method
- It is designed to test hypotheses and rigorously monitor and assess outcomes.

Goals when testing medical treatments:

- Efficacy: learn whether the treatment has a beneficial therapeutic effect
- Safety: learn whether the treatment is safe enough

Clinical trial: protocol and types

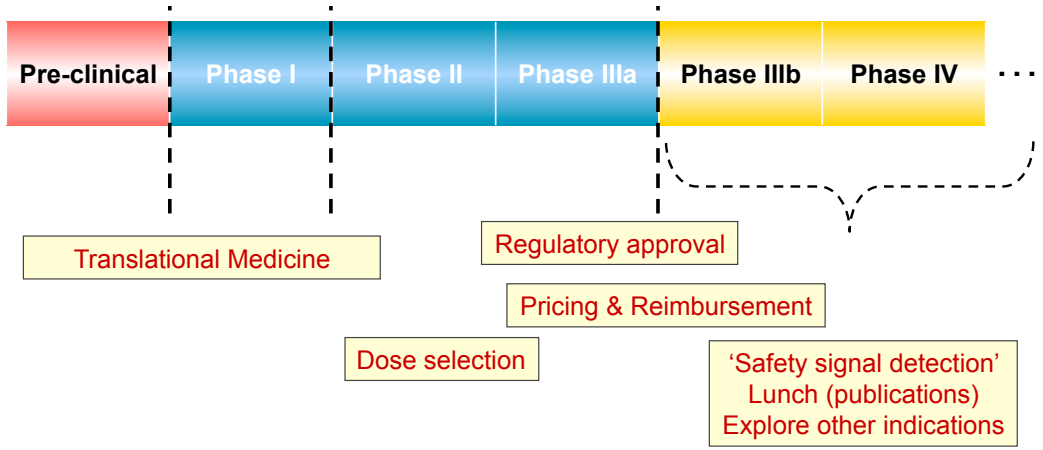
Clinical trial protocol: It is the trial's 'operating manual' specifying the design and objectives of the trial and ensuring that all researchers perform the trial in the same way on similar patients and that the data is comparable across all patients.

NIH classifies trials into five different types:

1. Prevention trials: better ways to prevent disease in people who have never had the disease
2. Screening trials: best way to detect certain diseases
3. Diagnostic trials: better procedures for diagnosing a particular disease
4. Treatment trials: test experimental treatments: new combinations of drugs, new therapeutical approaches
5. Quality of life: explore ways to improve comfort and quality of life for individuals with a chronic illness.

Phases of a clinical trial

0. Phase 0: Pharmacodynamics and pharmacokinetics in humans. Number of subjects: 10 to 15 to gather preliminary data
1. Phase 1: Screening for safety. Number of subjects: 20-80 to evaluate safety, determine safe dosage ranges, and identify side effects.
2. Phase 2: Efficacy of the drug. Number of subjects: 100-300 to see if it is effective and to evaluate its safety.
3. Phase 3: Final confirmation of safety and efficacy. Number of subjects: 1,000-3,000 to confirm effectiveness, monitor side effects and collect information.
4. Phase 4: Postmarketing studies. Additional information under "normal" use



Decision for primary endpoint

Primary endpoint

Variable (outcome) measuring the clinical evidence: it is a key decision for the study because the efficacy of the new drug and sample size computation will based on the primary endpoint

Improvements in medical management have led to:

- Decline in mortality and morbidity for several common disorders $\Rightarrow$ Low event rates
- Decline in the incidence of clinically relevant outcomes $\Rightarrow$ Reduction in the number of relevant events
- Improved standard of care $\Rightarrow$ Lower effect sizes

HENCE **relevant endpoints** are observed **less often** and the effect of treatment is **diminished**.

COMPOSITE ENDPOINTS might be a solution

DEFINITION OF COMPOSITE ENDPOINTS

Composite Outcome

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

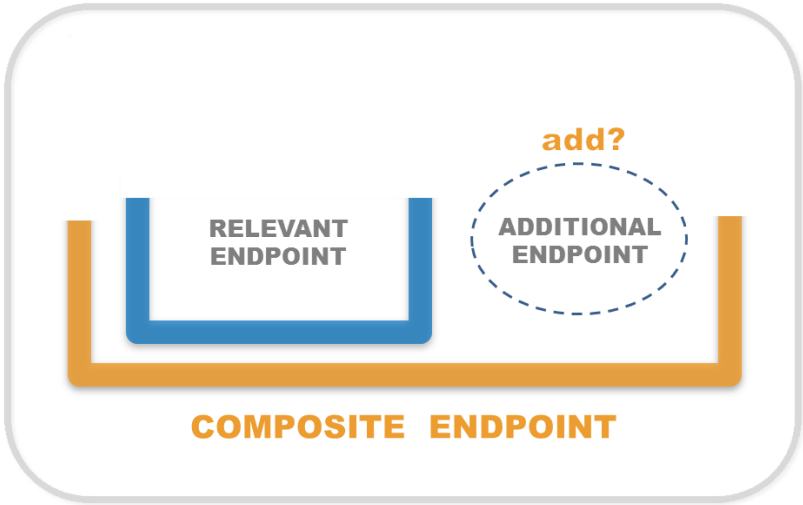
Composite Endpoint (CE)

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

ICH E9 guideline: *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)].

[(1)] ICH: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use

WHY TO USE COMPOSITE ENDPOINTS?



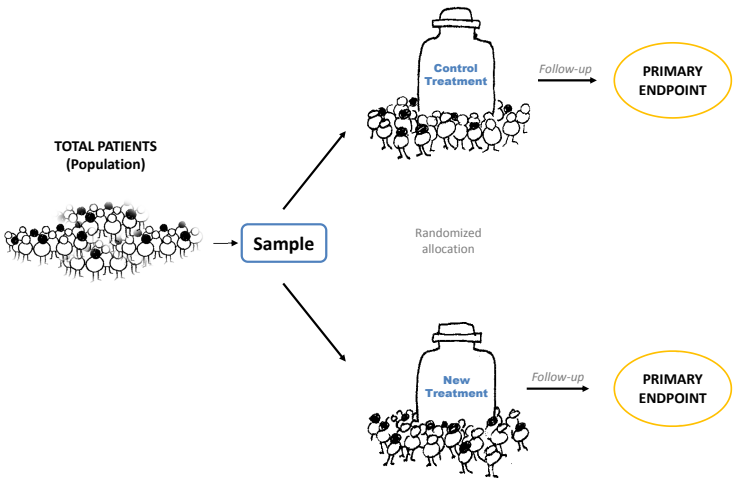
1. To get a better description of the disease process
2. To estimate the net clinical benefit of a therapy
3. To get higher event rates
4. To avoid adjustment for multiple comparisons
5. To avoid competing risks
6. TO IMPROVE STATISTICAL EFFICIENCY BY
   - needing smaller sample sizes
   - shorter follow-up times

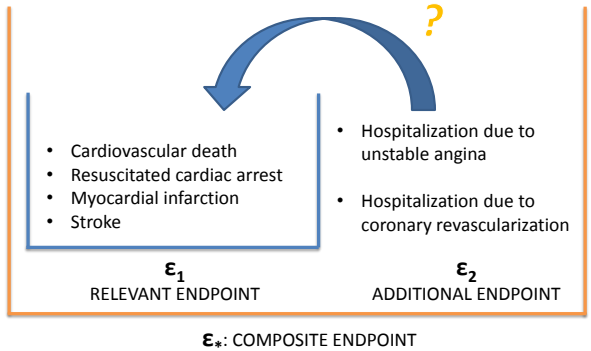
SOME AREAS WHERE CE ARE USED?:

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. **CARDIOVASCULAR DISEASE STUDIES**
   - T_1 time to the first event between CV death, MI, stroke ($\mathcal{E}_1$)
   - T_2 time to Hospitalization ($\mathcal{E}_2$)
   - T_* time to MACE: Major Adverse Cardiovascular Events ($\mathcal{E}_*$)
3. **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 ($\mathcal{E}_*$)
4. **NEUROLOGICAL STUDIES, PROSTATE PREVENTION STUDIES, ...**

SETTING THE PROBLEM

RCT for comparing the efficacy of new treatment $X = 1$ versus standard of care $X = 0$



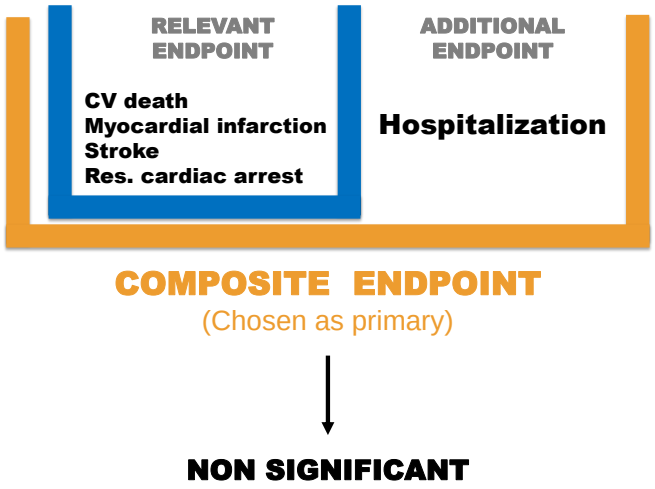


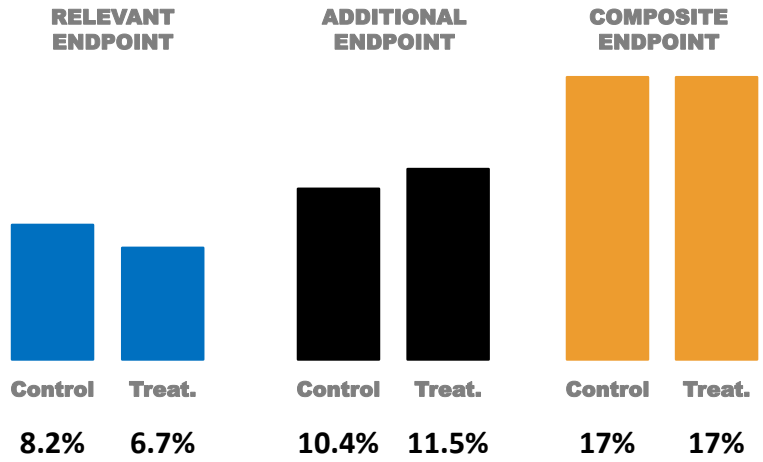
- RELEVANT ENDPOINT
  T_1 =time to $\mathcal{E}_1$: time to first between CV death; cardiac arrest; MI; stroke
- ADDITIONAL ENDPOINT
  T_2 =time to $\mathcal{E}_2$: time to hosp
- COMPOSITE ENDPOINT
  T_* =time to $\mathcal{E}_1 \cup \mathcal{E}_2$: time to MACE.

ARISE TRIAL[2]

Acute coronary syndrome patients randomized to receive:

0. Standard of care ($n = 3066$)
   - 252 events $\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4$
   - 277 events $\mathcal{E}_5 \cup \mathcal{E}_6$
   - 529 events $\mathcal{E}_* = \bigcup_{i=1}^{6} \mathcal{E}_i$
1. SOC+Succinobucol ($n = 3078$)
   - 207 events $\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4$
   - 323 events $\mathcal{E}_5 \cup \mathcal{E}_6$
   - 530 events $\mathcal{E}_* = \bigcup_{i=1}^{6} \mathcal{E}_i$





1. The CV hospital admission component MASKED the mortality effect
2. Succinobucol might have shown a beneficial effect ($p = 0.029$) on $\mathcal{E}_1$

[2] Tardif JC et al. Effects of succinobucol (AGI-1067) after an acute coronary syndrome: a randomised, double-blind, placebo-controlled trial(2008). *The Lancet*. 371, Issue 9626, 1761-1768

TESTING THE TREATMENT EFFECT

If Primary Endpoint is based on T_1 =time to $\mathcal{E}_1$

- H_0 : Treatment has **No EFFECT** on time to RELEVANT ENDPOINT
- H_1: **EFFECT** of treatment on time to RELEVANT ENDPOINT

If Primary Endpoint is based on $T_* = \min(T_1, T_2)$, the composite of $\mathcal{E}_1$ and $\mathcal{E}_2$ where T_2 =time to $\mathcal{E}_2$ is an additional endpoint.

- H_0^* :Treatment has **No EFFECT** on time to COMPOSITE ENDPOINT
- H_1^*: **EFFECT** of treatment on time to COMPOSITE ENDPOINT

H_0 and H_0^* ARE NOT EQUIVALENT HYPOTHESES !!!!! $\Rightarrow$ MORE LATER and ONGOING RESEARCH

Logrank test Z for H_0 vs H_1 via T_1

H_0: $\mathrm{HR}(t) = \frac{\lambda_1^{(1)}(t)}{\lambda_1^{(0)}(t)} = 1$ where $\lambda_1^{(j)}(t)$ is the hazard function for $T_1|X = j$

Log-rank test is used to compare the survival distributions of two samples.

It compares estimates of the hazard functions of the two groups at each observed event time.

It takes into account the risk set $R^{(j)}(t)$: patients at risk in each group where there is an event.

ASYMPTOTIC BEHAVIOUR of Z:

- Logrank $Z \sim N(0,1)$ under H_0
- Logrank $Z \sim N(\mu, 1)$ under fixed alternatives H_1: $\mathrm{HR}(t) = HR_1$
- Logrank $Z \sim N(\mu_1, 1)$ under a sequence of contiguous alternatives closer to H_0, that is, $H_{1,n}$: $\log \mathrm{HR}_n(t) = \log\left(\frac{\lambda_{1,n}^{(1)}(t)}{\lambda_1^{(0)}(t)}\right) = \frac{g(t)}{\sqrt{n}}$

Is death one of the components of T_1 or T_2?

If $\mathcal{E}_1$ and $\mathcal{E}_2$ are the two components of $\mathcal{E}_*$, if one contains death, it precludes the observation of the other and is a competing cause

- Add. endpoint $\mathcal{E}_2$ not including death (T_1 competing cause for T_2) $\Rightarrow$
  Case 1: $\mathcal{E}_1$ does not contain death
  Case 3: $\mathcal{E}_1$ contains death
  - T_1 censored by C (end-of-study censoring)
  - T_2 censored by $\min(C, T_1)$
  - T_* censored by C
- Add. endpoint $\mathcal{E}_2$ contains death (T_2 competing cause for T_1) $\Rightarrow$
  Case 2: $\mathcal{E}_1$ does not contain death
  Case 4: $\mathcal{E}_1$ contains death
  - T_1 censored by $\min(C, T_2)$
  - T_2 censored by C
  - T_* censored by C

Each censoring case has to be worked separately because involve different marginal or cause-specific hazards

Marginal and cause specific hazard functions

- Censoring cases 1 and 3: $\mathcal{E}_2$ does not contain death

  Marginal hazard function for T_1 in group j can be identified from observable data and can be used for Hypothesis testing

  $\lambda_1^{(j)}(t) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \mathrm{P}[t \leq T_1^{(j)} < t + \Delta t \mid T_1^{(j)} \geq t]$
- Censoring cases 2 and 4: $\mathcal{E}_2$ contains death

  $\lambda_1^{(j)}(t)$ cannot be identified from observable data and cannot be used for Hypothesis testing. Inference has to be based on

  Cause-specific hazard for T_1 in group j

  $$\lambda_{C1}^{(j)}(t) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \mathrm{Prob}\{t \leq T_1^{(j)} < t + \Delta t, T_1^{(j)} < T_2^{(j)} | T_*^{(j)} \geq t\}$$
- Marginal hazard function for $T_* = \min\{T_1, T_2\}$ in group j for all 4 censoring cases

  $\lambda_*^{(j)}(t) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} \mathrm{P}[t \leq T_*^{(j)} < t + \Delta t \mid T_*^{(j)} \geq t]$

  Needs the law of (T_1, T_2) to obtain the law of T_*

Asymptotic behaviour of Z under sequence of contiguous alternatives to H_0

Z is asymptotically $N(\mu, 1)$ [3] where

$$\frac{\mu}{\sqrt{n}} = \frac{\int_0^\infty p(t)[1-p(t)]\log\left\{\mathrm{HR}_n(t)\right\}V(t)dt}{\sqrt{\int_0^\infty p(t)[1-p(t)]V(t)dt}}$$

- U is the observed outcome
  - Cases 1 and 3: $U = \min\{T_1, C\} \Rightarrow P_{H_0}(U \geq t) = P_{H_0}(T_1 > t, C \geq t)$
  - Cases 2 and 4: $U = \min\{T_1, T_2, C\} \Rightarrow$ $P_{H_0}(U \geq t) = P_{H_0}(T_1 > t, T_2 > t, C \geq t)$
- $p(t) = P_{H_0}(X = 1 | U \geq t)$ is null prob. someone at risk at t is in gr. 1
- $V(t) = P_{H_0}(U \geq t)\lambda_1^{(0)}(t)dt = P_{H_0}(T_1 > t, C \geq t)\lambda_1^{(0)}(t)dt$ null sub-density function of observing a T_1 event at time t.

[3] Lagakos S.W. and Schoenfeld, D. Properties of Proportional-Hazards Score Tests under Misspecified Regression Models (1984). *Biometrics*, **40**, 1037–1048.

Testing H_0^* vs H_1^*: Logrank test Z_* for T_* (the same for 4 censoring cases)

- $\lambda_*^{(0)}(t)$, $\lambda_*^{(1)}(t)$ hazards for T_*
- $H_0^* : \mathrm{HR}_*(t) = \frac{\lambda_*^{(1)}(t)}{\lambda_*^{(0)}(t)} = 1 \Leftrightarrow$ NO EFFECT on T_*
- $Z_* \sim N(0,1)$ under H_0^*
- $Z_* \sim N(\mu_*, 1)$ under $H_{*,n} := \log\left(\frac{\lambda_{*,n}^{(1)}(t)}{\lambda_*^{(0)}(t)}\right) = \frac{g_*(t)}{\sqrt{n}}$

$$\frac{\mu_*}{\sqrt{n}} = \frac{\int_0^\infty p_*(t)[1-p_*(t)] \log\left\{\frac{\lambda_{*,n}^{(1)}(t)}{\lambda_*^{(0)}(t)}\right\} V_*(t)dt}{\sqrt{\int_0^\infty p_*(t)[1-p_*(t)] V_*(t)dt}}$$

  - **We need the law of (T_1, T_2). We'll discuss later**
  - $p_*(t) = P_{H_0^*}(X = 1 | U_* \geq t)$ null prob. someone at risk at t is in gr. 1
  - $V_*(t) = P_{H_0}(U_* \geq t)\lambda_*^{(0)}(t)dt = P_{H_0^*}(T_* > t, C \geq t)\lambda_*^{(0)}(t)dt$ null sub-density function of observing a T_* event at time t

Asymptotic Relative Efficiency (ARE)

TO ASSESS RELATIVE EFFICIENCY BETWEEN USING $\mathcal{E}_1$ VERSUS USING THE COMPOSITE $\mathcal{E}^* = \mathcal{E}_1 \cup \mathcal{E}_2$ [4]

- $Z \sim N(\mu, 1)$
- $Z_* \sim N(\mu_*, 1)$
- $$\mathrm{ARE}\,(Z_*, Z) = \left(\frac{\mu_*}{\mu}\right)^2$$

We will assume for the purpose of implementation:

- Equal number of subjects in the two treatment groups.
- End-of-study censoring C at time τ: only noninf. cens. cause
- C identical across groups.
- Cases 1 and 3: $\frac{\lambda_1^{(1)}(t)}{\lambda_1^{(0)}(t)} = \mathrm{HR}_1$ and $\frac{\lambda_2^{(1)}(t)}{\lambda_2^{(0)}(t)} = \mathrm{HR}_2$: Constant treatment HR for T_1 and T_2
- Cases 2 and 4: Constant treatment cause-specific HR for T_1 and T_2

[4] Gómez G. and Lagakos S.W. Statistical considerations when using a composite endpoint for comparing treatment groups (2013). Statistics in Medicine, 32, 719–738.

Pitman's Interpretation of ARE [5]

The efficacy of the treatment is set by means of two sets of hypotheses:

$H_0 : \mathrm{HR}_1(t) = 1$ versus $H_1 : \mathrm{HR}_1(t) = h_R < 1$ to be conducted by means of logrank test Z based on $\mathcal{E}_1$ and

$H_0^* : \mathrm{HR}_*(t) = 1$ versus $H_1^* : \mathrm{HR}_*(t) < 1$ to be conducted by means of logrank test Z_* based on $\mathcal{E}_*$.

Let n_1 and n_* be the sample sizes required for Z and Z_* to have power $1-\beta$ at level α.

Given $0 < \alpha < 1-\beta < 1$, $\boxed{\mathrm{ARE} \approx \frac{n_1}{n_*}}$

Open questions here!!!

Criterion for Decision

$\mathrm{ARE}(Z_*, Z) > 1 \Rightarrow T_*$ **more efficient than** $T_1 \Rightarrow$ Use composite endpoint

[5] Gómez G. and Gómez-Mateu M. (2014). The Asymptotic Relative Efficiency and the ratio of sample sizes when testing two different null hypotheses. SORT, 38, 73–88.

Computing ARE for cases 1 and 3

$$\mathrm{ARE}\,(Z_*, Z) = \left(\frac{\mu_*}{\mu}\right)^2 = \frac{\left(\int_0^1 \log\left\{\frac{\lambda_*^{(1)}(t)}{\lambda_*^{(0)}(t)}\right\} f_*^{(0)}(t)dt\right)^2}{(\log \mathrm{HR}_1)^2\left(\int_0^1 f_*^{(0)}(t)dt\right)\left(\int_0^1 f_1^{(0)}(t)dt\right)}$$

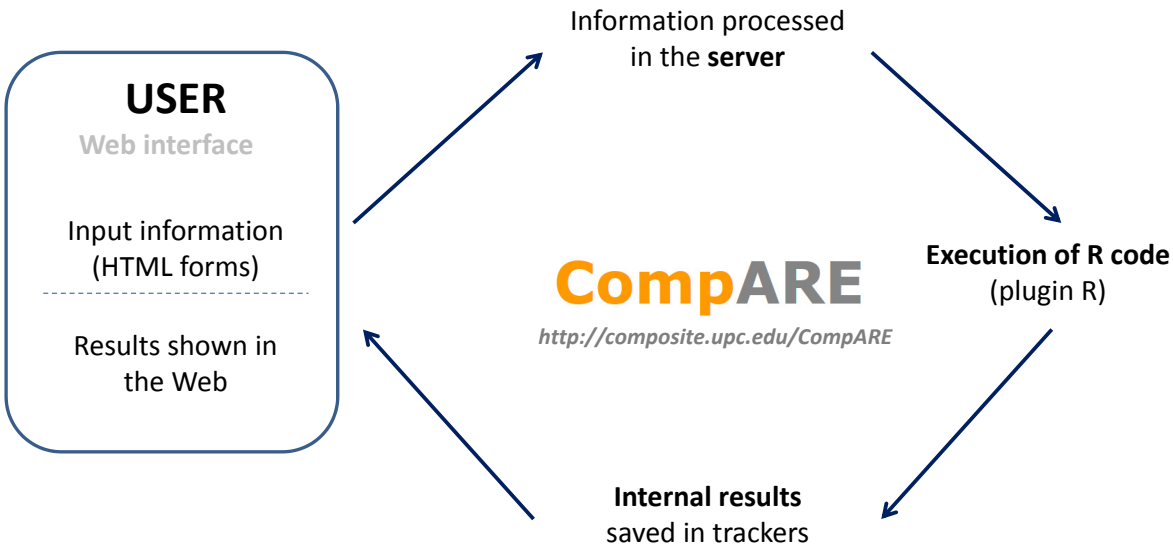
- **Depends on the relevant endpoint T_1 via**
  - Marginal density $f_1^{(0)}(t)$ (assumed Weibull)
  - $p_1 =$ Probability of observing T_1 in group 0
  - $\mathrm{HR}_1 = \frac{\lambda_1^{(1)}(t)}{\lambda_1^{(0)}(t)}$ relative treatment effect on $\mathcal{E}_1$
- **Depends on the joint distribution of (T_1, T_2) via**:
  - Copula binding the marginal densities (both assumed Weibull). **Technicalities later**
  - ρ: Spearman's rank correlation between $T_1^{(0)}$ and $T_2^{(0)}$ (assumed equal for both groups)
  - $p_2 =$ Probability of observing T_2 in group 0
  - $\mathrm{HR}_2 = \frac{\lambda_2^{(1)}(t)}{\lambda_2^{(0)}(t)}$ relative treatment effect on $\mathcal{E}_2$

ARE method implemented in CompARE

1. Set values for $p_1, p_2, \mathrm{HR}_1, \mathrm{HR}_2, \rho$
2. Assume **Weibull** $(b_1^{(j)}, \beta_1^{(j)})$ for T_1 and **Weibull** $(b_2^{(j)}, \beta_2^{(j)})$ for T_2
3. Assume $\beta_k = \beta_k^{(0)} = \beta_k^{(1)}$ (for $k = 1, 2$) so that the proportionality of the hazards holds
4. Set values for shape parameters β_1 and β_2
5. Compute scale parameters as
   1. $b_1^{(0)}(p_1, \beta_1) = \frac{1}{(-\log(1-p_1))^{1/\beta_1}}$
   2.
      1. $b_2^{(0)}(p_2, \beta_2) = \frac{1}{(-\log(1-p_2))^{1/\beta_2}}$ if $\mathcal{E}_1$ does not include a terminating event
      2. $b_2^{(0)}(p_1, p_2, \rho, \beta_1, \beta_2)$ is the solution of $p_2 = \int_0^1 \int_v^\infty f_{(1,2)}^{(0)}(u, v; \theta) du dv$ if $\mathcal{E}_1$ includes a terminating event
   3. $b_k^{(1)}(b_k^{(0)}, \beta_k, \mathrm{HR}_k) = \frac{b_k^{(0)}}{\mathrm{HR_k}^{1/\beta_k}}$ for $k = 1, 2$
6. Get association parameter θ from Spearman's ρ
7. Compute **Copula** $C(S_{T_1}(t_1), S_{T_2}(t_2); \theta)$ for both groups ($X = 0$ and $X = 1$) using equal θ for both groups
8. Get $\mathrm{ARE}(Z_*, Z)$ as function of $p_1, p_2, \mathrm{HR}_1, \mathrm{HR}_2, \rho$

CompARE web platform

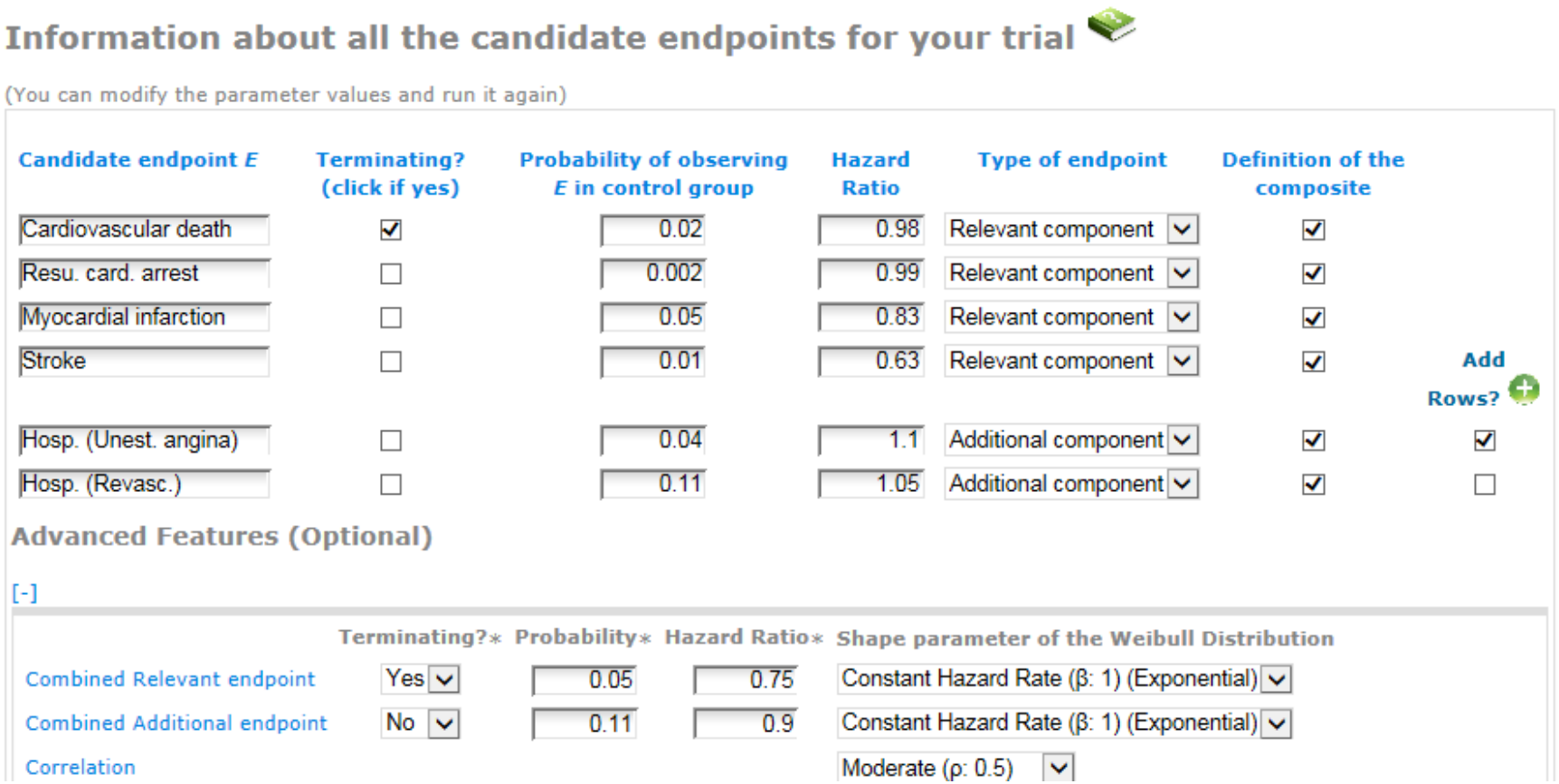
- Free and easy to use
- Knowledge of R not needed
- Accessible anywhere (laptop/mobile/tablet)
- Compatible with any operating system and browser
- Complete users' guide documentation



Software used to built the Interface

- **Tiki=** Tightly Integrated Knowledge Infrastructure. *Free and Open Source Web Application with built-in features.*
- **Wiki**: *Website which allows its users to add, modify, or delete its content via a web browser usually using a simplified markup language or a rich-text editor*

Treating patients with succinobucol. Analysis with CompARE (Tardif *et al.* Lancet 2008)



Information about all the candidate endpoints for your trial

(You can modify the parameter values and run it again)

Candidate endpoint *E*	Terminating? (click if yes)	Probability of observing *E* in control group	Hazard Ratio	Type of endpoint	Definition of the composite	
Cardiovascular death	☑	0.02	0.98	Relevant component	☑	
Resu. card. arrest	☐	0.002	0.99	Relevant component	☑	
Myocardial infarction	☐	0.05	0.83	Relevant component	☑	
Stroke	☐	0.01	0.63	Relevant component	☑	Add Rows?
Hosp. (Unest. angina)	☐	0.04	1.1	Additional component	☑	☑
Hosp. (Revasc.)	☐	0.11	1.05	Additional component	☑	☐

Advanced Features (Optional)

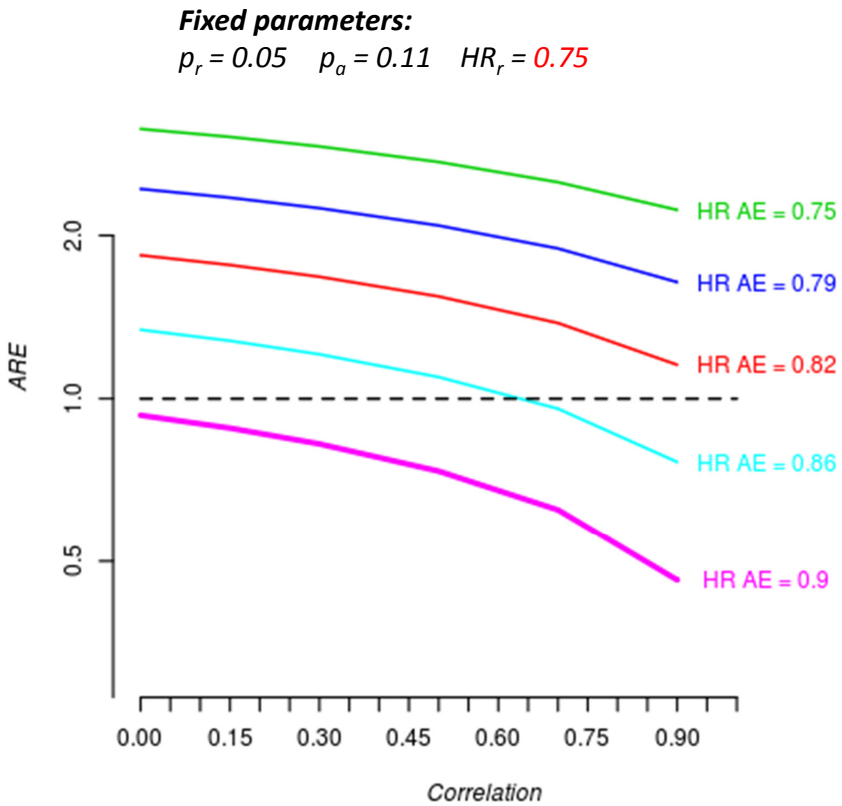
[-]

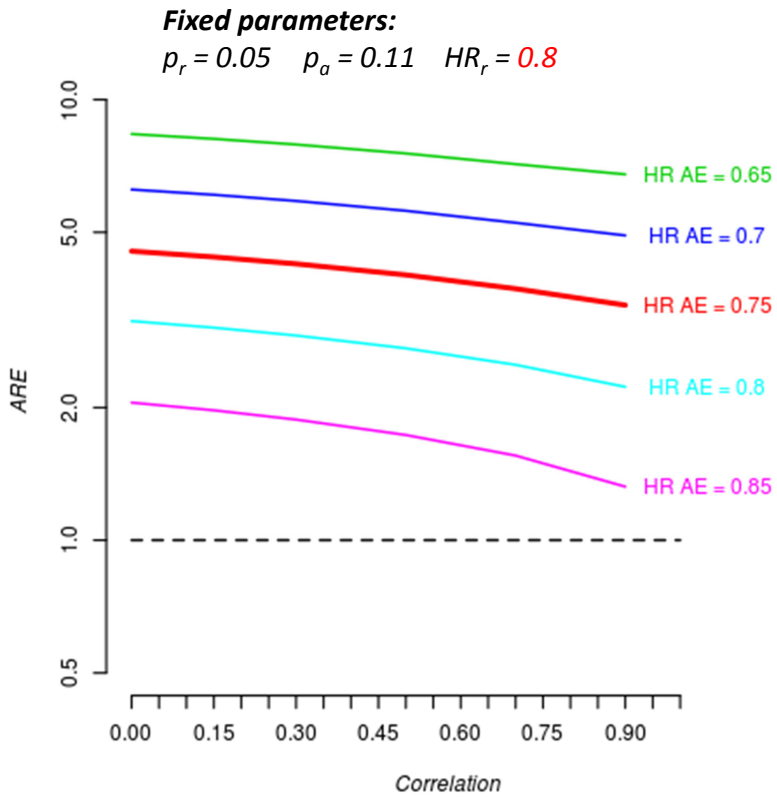
	Terminating?∗	Probability∗	Hazard Ratio∗	Shape parameter of the Weibull Distribution
Combined Relevant endpoint	Yes	0.05	0.75	Constant Hazard Rate (β: 1) (Exponential)
Combined Additional endpoint	No	0.11	0.9	Constant Hazard Rate (β: 1) (Exponential)
Correlation				Moderate (ρ: 0.5)

CompARE GRAPHICAL RESULTS

$\mathrm{HR}_2 = 1.05 \Rightarrow \mathrm{ARE}(T_* \mathrm{vs} T_1) < 1, \ \forall \ \rho(T_1, T_2) \Rightarrow \mathcal{E}_1$ should have been used.

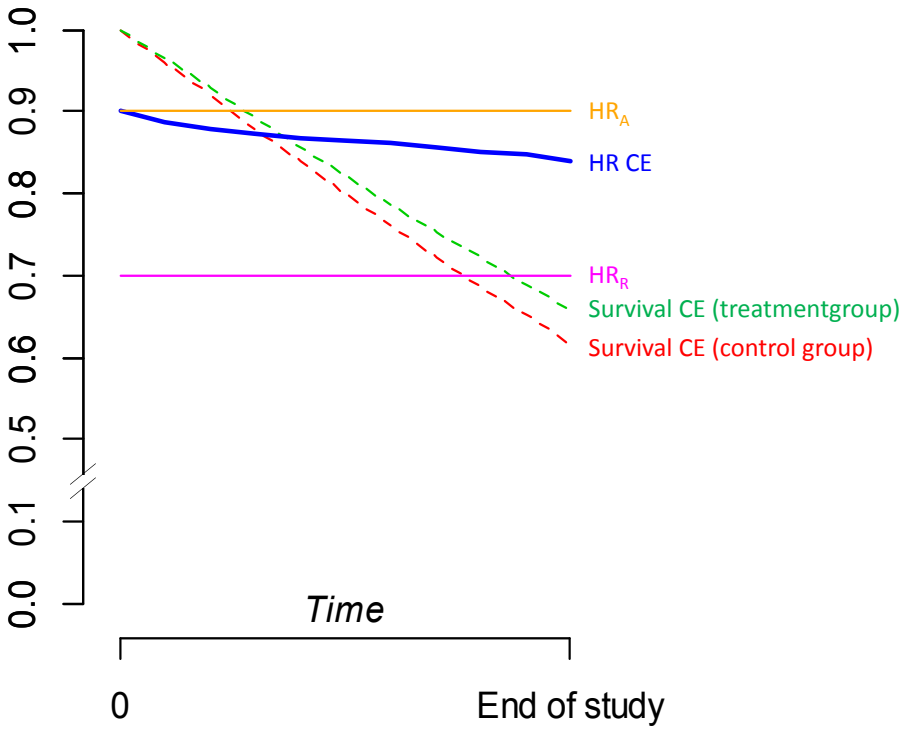
$\mathcal{E}_*$ would have been justified if $\mathrm{HR}_2 \leq 0.88$ [(6)]





[(6)] Gómez G, Gómez-Mateu M. Comments on "Use of composite endpoints in clinical trials" by Abdul J. Sankoh, Haihong Li and Ralph B. D'Agostino, Sr (2015). *Statistics in Medicine*. (*in print*).

CompARE OTHER OUTPUTS AND TABLES



ARE results depending on different correlation values and Hazard Ratios

Fixed parameters:	
Probability RE (Control group)	0.15
Probability AE (Control group)	0.3
Hazard Ratio RE	0.7
Distribution RE	Increasing Hazard Rate
Distribution AE	Constant Hazard Rate (exponential)

Hazard Ratio AE	Correlation	ARE	Recommendation
0.9	0	0.64	Use RE
0.9	0.15	0.56	Use RE
0.9	0.3	0.49	Use RE
0.9	0.5	0.39	Use RE
0.9	**0.7**	**0.3**	**Use RE**
0.9	0.9	0.21	Use RE
0.7	0	2.78	Use CE
0.7	0.15	2.59	Use CE
0.7	0.3	2.4	Use CE
0.7	0.5	2.18	Use CE
0.7	0.7	1.99	Use CE
0.7	0.9	1.9	Use CE

RELATED ISSUES (TIME PERMITTED...)

- Copulas
- Non Proportional Hazards
- Sample size under NPH

Copula for the Law of (T_1, T_2)

A copula is a bivariate distribution on uniform random variables:

- marginal distributions $F_1(t)$, $F_2(t)$ are binded to form the joint $F(t_1, t_2; \theta) = C(F_1(t_1), F_2(t_2); \theta)$
- θ parameterises the dependence between the margins
- Different types of dependence can be represented

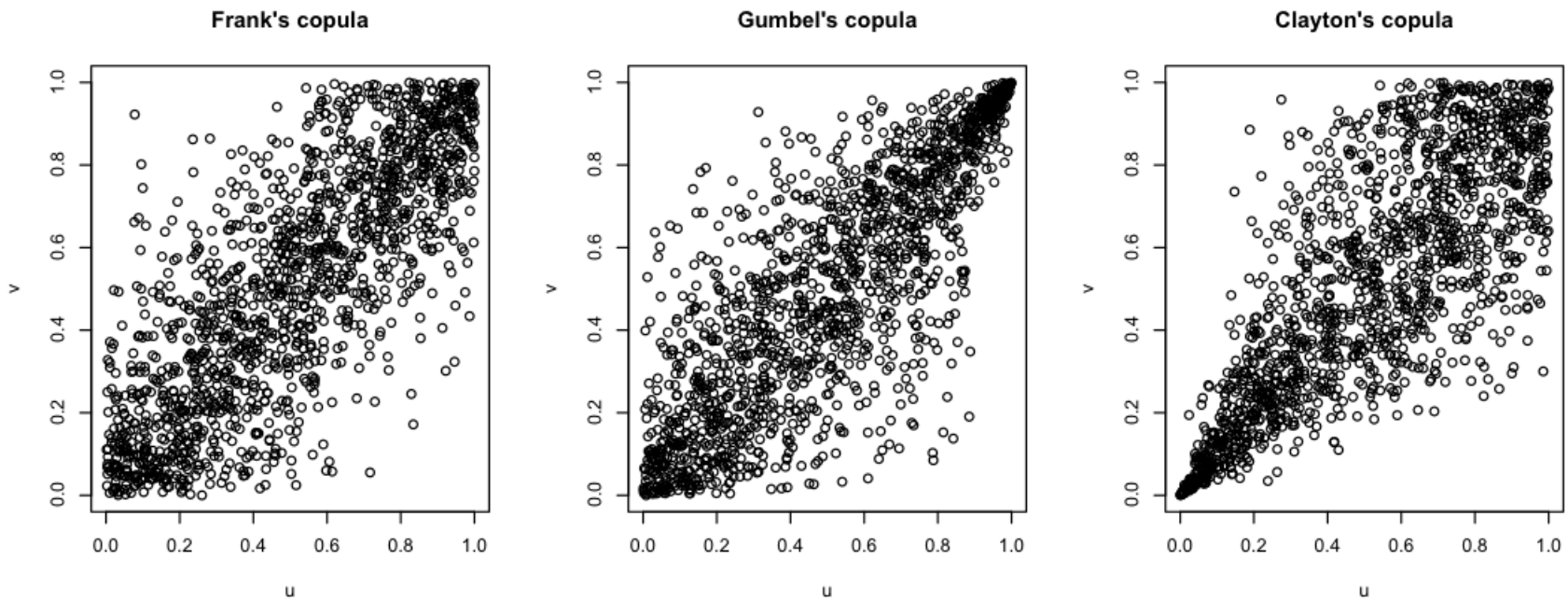


Figure: Common bivariate copulas

Frank's copula for (T_1, T_2)

1. Frank's copula function:

$$C(u_1, u_2; \theta) = -\theta^{-1} \log \left\{ 1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right\}.$$

   - θ, 1-1 function of Spearman's ρ, accounts for the dependency between T_1 and T_2

2. Joint density function for (T_1, T_2):

$$f_{(T_1,T_2)}(t_1, t_2; \theta) = \frac{\theta}{1 - e^{-\theta}} \frac{e^{-\theta(S_{T_1}(t_1)+S_{T_2}(t_2))}}{e^{-2\theta C(t_1,t_2;\theta)}} [f_{T_1}(t_1)][f_{T_2}(t_2)]$$

3. Density function of $T_* = \min\{T_1, T_2\}$

$$f_*(t; \theta) = \frac{e^{-\theta S_{T_1}(t)}(e^{-\theta S_{T_2}(t)} - 1) f_{T_1}(t)}{e^{-\theta C(S_{T_1}(t),S_{T_2}(t);\theta)}(e^{-\theta} - 1)} + \frac{e^{-\theta S_{T_2}(t)}(e^{-\theta S_{T_1}(t)} - 1) f_{T_2}(t)}{e^{-\theta C(S_{T_1}(t),S_{T_2}(t);\theta)}(e^{-\theta} - 1)}$$

Robustness w.r.t. choice of the copula

ARE implemented for 11 different copulas[7]

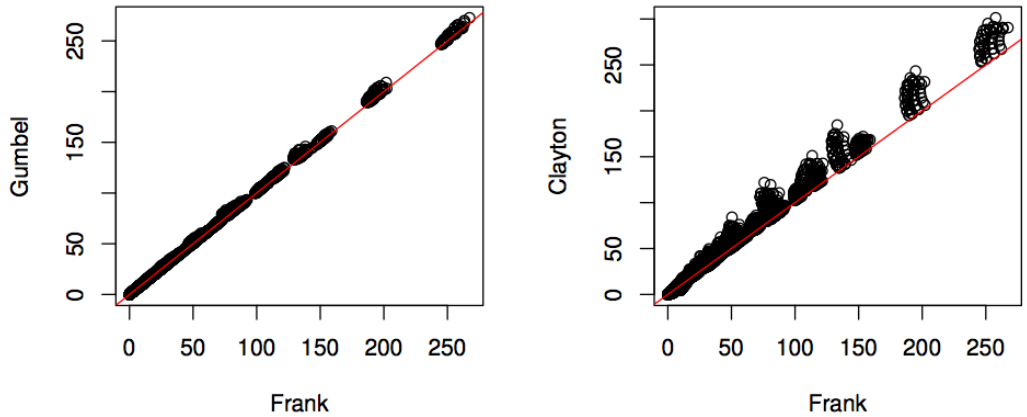


Figure: Pairwise ARE correlations based on 72576 simulated situations

ARE Comparisons	**Pearson's ρ**	**Spearman's ρ**	**Kendall's τ**
Frank - Gumbel	0.99987	0.99946	0.98229
Frank - Clayton	0.99701	0.99150	0.92735

Degree of agreement in recommendation **Frank - Gumbel** $\rightarrow$ 98.0%

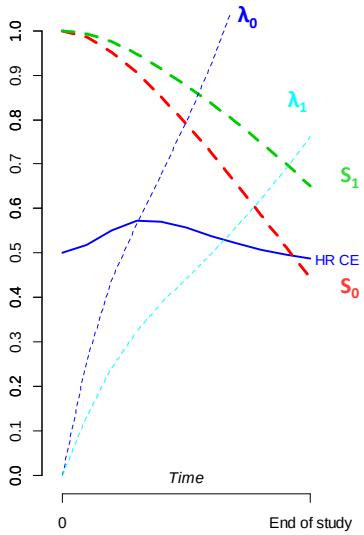
Degree of agreement in recommendation**Frank - Clayton** $\rightarrow$ 94.7%

Degree of agreement in recommendation **other pairs** $\rightarrow > 90\%$

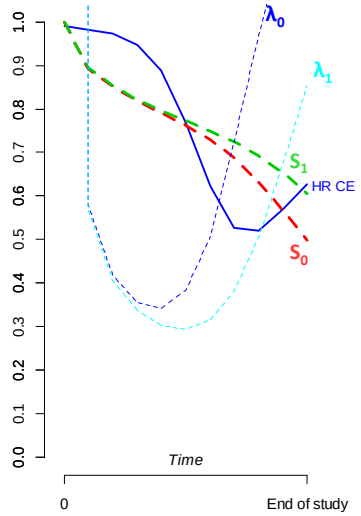
[7] Plana, O. and Gómez G. Selecting the primary endpoint in a randomized clinical trial. The ARE method. J. Biopharmaceutical Statist. Online Sept 2015

Non Proportional hazards

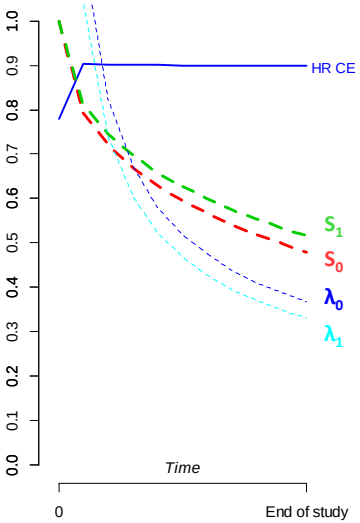
Diverging



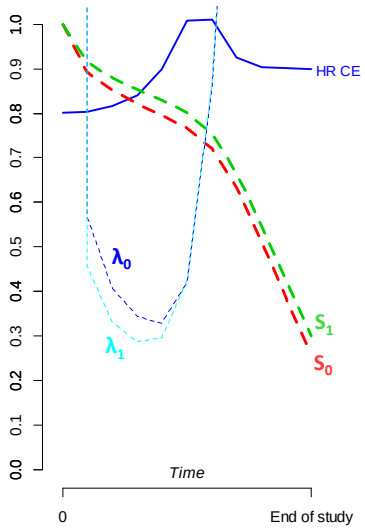
Late emerging difference



Parallel after t0



Transient difference



Most of the analyses use HR constant, however ...

$T_{CE} = \min\{T_R, T_A\}$, $\mathrm{HR_R}$ and $\mathrm{HR_A}$ constant $\Longrightarrow$ $\mathrm{HR_{CE}}(t)$ not necessarily constant

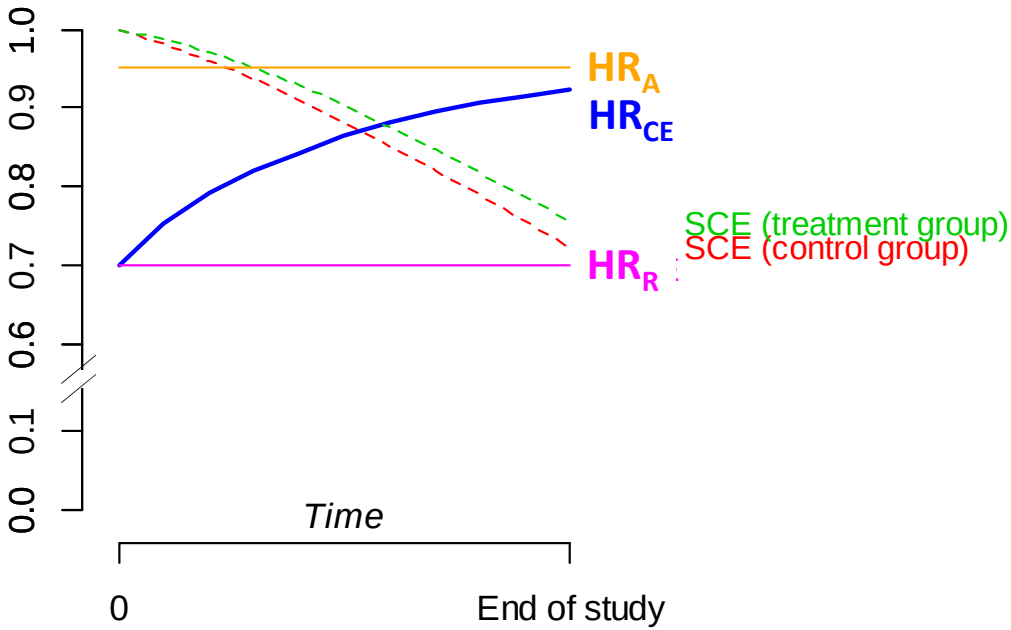
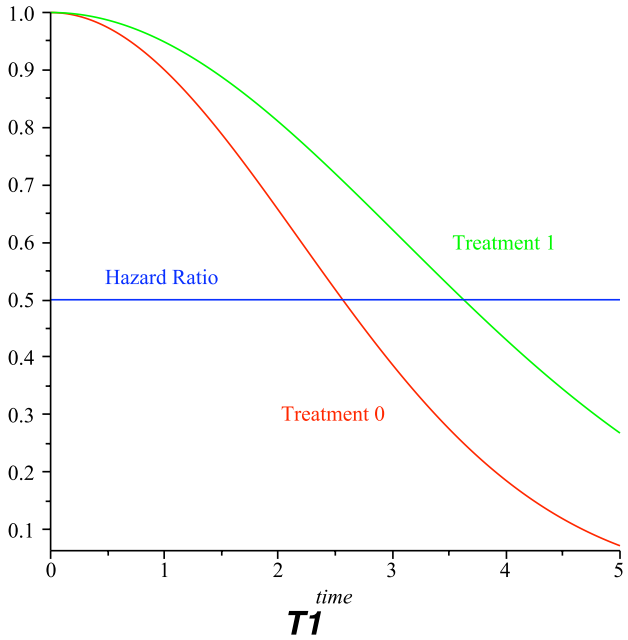
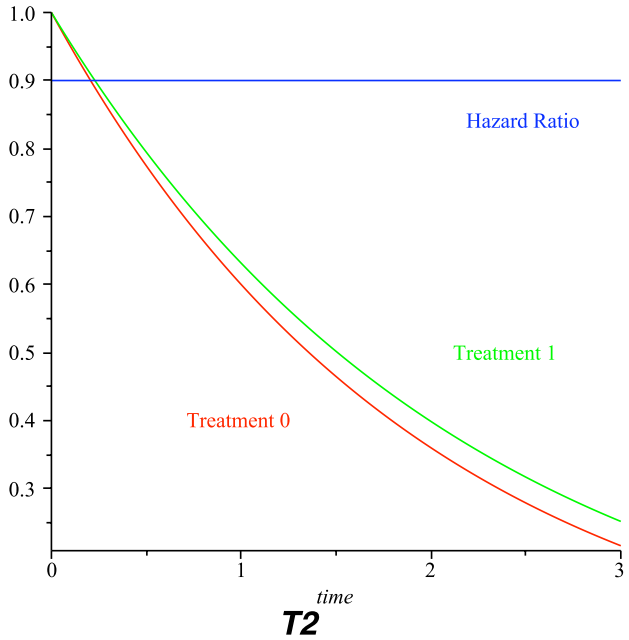


Figure: T_R (death) $\sim$ Weibull: $\beta_R = 1$,$p_R = 0.15$, $HR_R = 0.7$.
T_A (not death) $\sim$ Weibull: $\beta_A = 2$, $p_A = 0.15$, $HR_A = 0.95$.
Spearman $\rho(T_R, T_A) = 0.5$

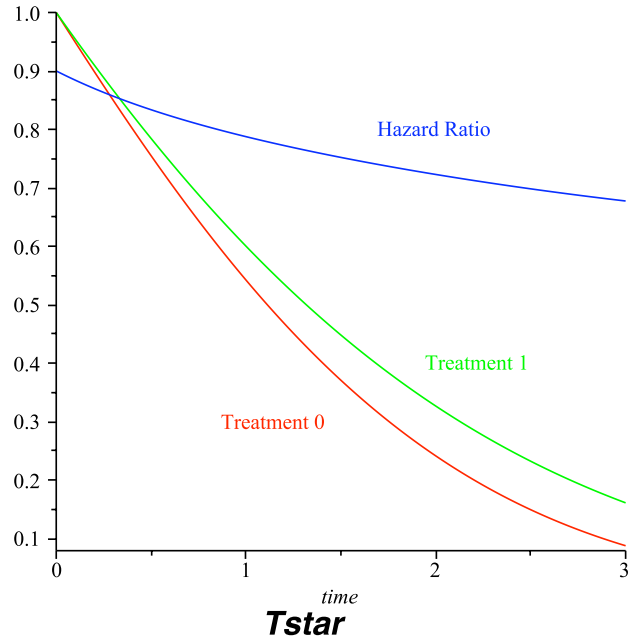
Behaviour of Hazard ratio for CE



T1



T2



Tstar

Proposition

$\lambda_*^{(1)}(t)$ and $\lambda_*^{(0)}(t)$ are proportional **if and only if** $\lambda_1^{(0)}(t)$ and $\lambda_2^{(0)}(t)$ are proportional. Then the hazard ratio $\mathrm{HR}_* = \dfrac{\lambda_*^{(1)}(t)}{\lambda_*^{(0)}(t)}$ is a linear combination of $\mathrm{HR}_1 = \dfrac{\lambda_1^{(1)}(t)}{\lambda_1^{(0)}(t)}$ and $\mathrm{HR}_2 = \dfrac{\lambda_2^{(1)}(t)}{\lambda_2^{(0)}(t)}$

Using ARE to get the sample size for $\mathcal{E}_{CE}$

$$\boxed{\text{ARE} \approx \frac{n_R}{n_{CE}}}$$

can be used to derive the sample size to detect $H_1^{(CE)} : \mathrm{HR_{CE}}(t) < 1$ based on formulas for the sample size n_R required to detect $H_1^{(R)} : \mathrm{HR_R}(t) = h_R < 1$

$$\boxed{n_{CE} \approx \frac{n_R}{ARE}}$$

Given $(\beta_1, \beta_2, p_R, p_A, \rho)$ and taking into acount if T_R or T_A include death, compute $\mathbf{A} = \mathrm{ARE}(\beta_R, \beta_A, p_R, p_A, h_R, h_A, \rho)$. For given α and $1-\beta$:

(A) If $\mathrm{ARE} \leq 1$, use T_R with sample size $n_R = \frac{4(z_\alpha + z_\beta)^2}{(\ln(h_R))^2 (\int_0^1 f_R^{(0)}(t)dt)}$

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

I $n_{CE} = \frac{4(z_\alpha + z_\beta)^2}{(\ln(h_{CE}))^2 (\int_0^1 f_{CE}^{(0)}(t)dt)}$ if $\mathrm{HR}_{CE}(t) = h_{CE}$ for all t,

II $n_{CE} = \frac{4(z_\alpha + z_\beta)^2}{\mathbf{A}(\ln(h_R))^2 (\int_0^1 f_R^{(0)}(t)dt)}$ if $\mathrm{HR}_{CE}(t) \leq h_{CE}$ for all t, not constant

Illustrating sample size for several degrees of dependence

p_R	HR_R
0,05	0,7

α	Power
0,05	0,8

p_A	HR_A	ρ	ARE	Red. SS CE (%)	N recommended
0,05	0,6	0,1	2,8	64,1	1.635
		0,5	2,5	59,6	1.841
		0,9	1,8	44,7	2.524
	0,7	0,1	1,9	47,7	2.386
		0,5	1,7	41,0	2.689
		0,9	1,2	19,2	3.683
	0,8	0,1	1,2	19,5	3.669
		0,5	1,1	9,0	4.149
		0,9	0,8	0	**4.560**
0,1	0,6	0,1	4,7	78,6	977
		0,5	4,1	75,6	1.112
		0,9	3,2	68,5	1.438
	0,7	0,1	2,8	64,6	1.613
		0,5	2,5	59,3	1.857
		0,9	1,8	45,5	2.486
	0,8	0,1	1,6	35,7	2.933
		0,5	1,3	24,3	3.452
		0,9	0,9	0	**4.560**

Ongoing Research and Technology Transfer

1. Behaviour of Hazard ratio in general. Can we characterize situations and find patterns for different scenarios?
2. Sample size calculation for non-proportional Hazard Ratios and based on alternative summary measures
3. Asymptotic Relative Efficiency for two non identical set of hypotheses. Theoretical insight
4. ARE method for binary outcomes
5. ARE for observational studies
6. Extensions of CompARE
   - to include sample size computations
   - Different copulas other than Frank's*
   - Combined Probabilities (in control group) and Hazard Ratio values
   - Computation of Sample Size based on ARE
   - Binary outcomes
   - Dynamic plots

SUMMARIZING

- Composite Endpoints have to be justified from a clinical point of view
- ARE: Conceptual framework as a tool to decide whether or not a CE should be used when comparing two treatment groups in a RCT
- **CompARE** to compute the ARE for time-to-event endpoints
- Hazard ratios for composite endpoints are almost always non proportional
- ARE method provides a framework to derive a formula for the required sample size
- Extending **CompARE** to sample size computation.
- ARE for binary CE and Observational studies
- Extending **CompARE** to binary CE.

Thanks to my coauthors



ARE method for Binary endpoints

▸ To ongoing research...

- $Y_{ij1} = \mathbf{1}\{RE\}$ with $p_{i1} = P(Y_{ij1} = 1)$ and
  $Y_{i1} = \sum_{j=1}^{N_i} Y_{ij1} \sim Bin(N_i, p_{i1})$
- $Y_{ij2} = \mathbf{1}\{AE\}$ with $p_{i2} = P(Y_{ij2} = 1)$ and
  $Y_{i2} = \sum_{j=1}^{N_i} Y_{ij2} \sim Bin(N_i, p_{i2})$
- $Y_{ij*} = \begin{cases} 1 & \text{if } Y_{ij1} + Y_{ij2} \geq 1 \\ 0 & \text{if } Y_{ij1} + Y_{ij2} = 0 \end{cases}$ with $p_{i*} = P(Y_{ij*} = 1)$ and
  $Y_{i*} = Y_{i1} + Y_{i2} \sim Bin(N_i, p_{i*})$, number responding to either RE or AE

Null Hypothesis

- $H_0 : p_{01} = p_{11} \Leftrightarrow \mathrm{OR}_1 = \frac{p_{11}/1-p_{11}}{p_{01}/1-p_{01}} = 1$
- $H_0^* : p_{0*} = p_{1*} \Leftrightarrow \mathrm{OR}_* = \frac{p_{1*}/1-p_{1*}}{p_{0*}/1-p_{0*}} = 1 \Leftrightarrow$
  $q_{01}q_{02} + \rho_0\sqrt{p_{01}p_{02}q_{01}q_{02}} = q_{11}q_{12} + \rho_1\sqrt{p_{11}p_{12}q_{11}q_{12}}$

Two Sample Binomial test statistics and ARE

Under $H_0 : p_{01} = p_{11}$

- $T_1 = \sqrt{N_0 + N_1}\, \frac{(N_0 Y_{11} - N_1 Y_{01})}{\sqrt{N_0 N_1 \tilde{p}_1 \tilde{q}_1}} \sim N(0,1)$

Under $H_{1,n}$**:** sequences of alternatives that converge to H_0

- $T_1 \sim N(\mu_1, 1)$
- $\mu_1^2 = \pi(1-\pi)(\log(\mathrm{OR}_1))^2\, p_{01} q_{01}$
- π is the probability of being allocated to control group

Under $H_0^* : p_{0*} = p_{1*}$

- $T_* = \sqrt{N_0 + N_1}\, \frac{(N_0 Y_{1*} - N_1 Y_{0*})}{\sqrt{N_0 N_1 \tilde{p}_* \tilde{q}_*}} \sim N(0,1)$

Under $H_{*,n}$**:** sequences of alternatives that converge to H_0^*

- $T_* \sim N(\mu_*, 1)$
- $\mu_*^2 = \pi(1-\pi)(\log(\mathrm{OR}_*))^2\, p_{0*} q_{0*}$

ARE: square of the ratio of the non-centrality means μ_1 and μ_*

$$ARE(T_1, T_*) = \left(\frac{\mu_*}{\mu_1}\right)^2 = \frac{(\log(\mathrm{OR}_*))^2\, p_{0*}(1-p_{0*})}{(\log(\mathrm{OR}_1))^2\, p_{01}(1-p_{01})}$$

Getting p_{i*} from p_{i1}, p_{i2} (subsequently OR_*)

Bahadur's theorem (1961)

The joint distribution between any pair of binary random variables is uniquely determined by the probabilities p_{i1}, p_{i2} and $\rho_i = Corr(Y_{ij1}, Y_{ij2})$,

$$P\left[Y_{ij1} = y_{ij1}, Y_{ij2} = y_{ij2}\right] = \prod_{k=1}^{2} \left(p_{ik}^{y_{ijk}} \cdot q_{ik}^{1-y_{ijk}}\right) \left(1 + \rho_i \cdot z_{ij1} \cdot z_{ij2}\right), \quad i = 0, 1$$

where $z_{ijk} = \frac{y_{ijk} - p_{ik}}{\sqrt{p_{ik} q_{ik}}}$ and $q_{ik} = 1 - p_{ik}$.

Correlation bounds (Sozu *et al.* (2010))

Given p_{i1} and p_{i2} $(i = 0, 1) \Rightarrow \rho_i = Corr(Y_{ij1}, Y_{ij2})$ is such that $-1 \leq M_I \leq \rho_i \leq m_i \leq 1$ where

$$M_i = \max\left\{-\sqrt{\frac{p_{i1}p_{i2}}{(1-p_{i1})(1-p_{i2})}}, -\sqrt{\frac{(1-p_{i1})(1-p_{i2})}{(p_{i1})(p_{i2})}}\right\}$$

and

ARE as a function of anticipatable parameters

The probability that an individual in group i has at least one response is

$$p_{i*} = 1 - P[Y_{ij*} = 0] = 1 - q_{i1}q_{i2} - \rho_i\sqrt{p_{i1}p_{i2}q_{i1}q_{i2}}$$

The Odds Ratio of having at least one response is

$$\begin{aligned} \mathrm{OR}_* &= \frac{(\mathrm{O}_{01}\mathrm{OR}_1 + 1)(\mathrm{O}_{02}\mathrm{OR}_2 + 1) - 1 - \rho_1\sqrt{\mathrm{O}_{01}\mathrm{OR}_1\mathrm{O}_{02}\mathrm{OR}_2}}{\frac{1}{q_{01}q_{02}} - 1 - \rho_0\sqrt{\mathrm{O}_{01}\mathrm{O}_{02}}} \\ &\cdot \frac{1 + \rho_0\sqrt{\mathrm{O}_{01}\mathrm{O}_{02}}}{1 + \rho_1\sqrt{\mathrm{O}_{01}\mathrm{OR}_1\mathrm{O}_{02}\mathrm{OR}_2}} \end{aligned}$$

where $\mathrm{O}_{01} = \frac{p_{01}}{1-p_{01}}$, $\mathrm{O}_{02} = \frac{p_{02}}{1-p_{02}}$ and $\mathrm{OR}_2 = \frac{p_{12}/1-p_{12}}{p_{02}/1-p_{02}}$.

We need to evaluate

$$ARE(T_*, T_1) = \frac{(\log(\mathrm{OR}_*))^2}{(\log(\mathrm{OR}_1))^2}\frac{p_{0*}(1-p_{0*})}{p_{01}(1-p_{01})}$$

1. Frequencies p_{01} and p_{02} of observing $\mathcal{E}_1$ and $\mathcal{E}_2$ in treatment group 0.