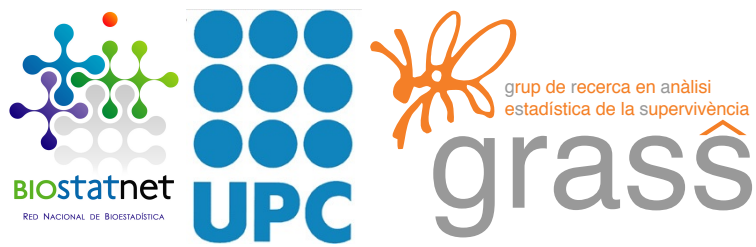
SELECTING THE PRIMARY ENDPOINT IN A RANDOMIZED CLINICAL TRIAL. How efficient is to add a new component to the primary endpoint?

Guadalupe Gómez
Universitat Politècnica de Catalunya, Barcelona, Spain

Department of Biostatistics
Harvard School of Public Health
November 21st, 2014



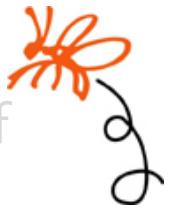
Randomized Clinical Trials

- **Goal: demonstrate the efficacy of a new drug**
- **Primary endpoint of a RCT: Outcome defined by the research question of interest**
- **Should be amenable to unbiased assessment and potentially influenced by the treatment**

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**



Randomized Clinical Trials

- **Goal: demonstrate the efficacy of a new drug**
- **Primary endpoint of a RCT: Outcome defined by the research question of interest**
- **Should be amenable to unbiased assessment and potentially influenced by the treatment**

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

Composite event: **union of a given set of events**

Composite endpoint (CE): **occurrence of first event, among a given set of events, after a certain period of follow-up.**

Why to use Composite Endpoints?:

1. A better description of the disease process
2. Gets higher event rates
3. Avoids adjustment for multiple comparisons
4. Avoids interpretational problems due to competing risks
5. Hopefully improves statistical efficiency and needs
   - smaller sample sizes
   - shorter follow-up times

Composite Endpoints

Composite event: **union of a given set of events**

Composite endpoint (CE): **occurrence of first event, among a given set of events, after a certain period of follow-up.**

Why to use Composite Endpoints?:

1. A better description of the disease process
2. Gets higher event rates
3. Avoids adjustment for multiple comparisons
4. Avoids interpretational problems due to competing risks
5. Hopefully improves statistical efficiency and needs
   - smaller sample sizes
   - shorter follow-up times

COMPOSITE ENDPOINTS IN SELECTED THERAPEUTIC AREAS

1. **CANCER CLINICAL TRIALS**
   DP: Disease progression
   OS: Overall survival
   PFS: Progression-free survival

2. **CARDIOVASCULAR DISEASE STUDIES**
   Cardiovascular death, myocardial infarction, stroke
   Hospitalization
   MACE: Major Adverse Cardiovascular Events

3. **HIV STUDIES**
   Virological failure
   Initiation of new treatment due to intolerance/toxicity
   TLOVR: Time to loss of virological response

COMPOSITE ENDPOINTS IN SELECTED THERAPEUTIC AREAS

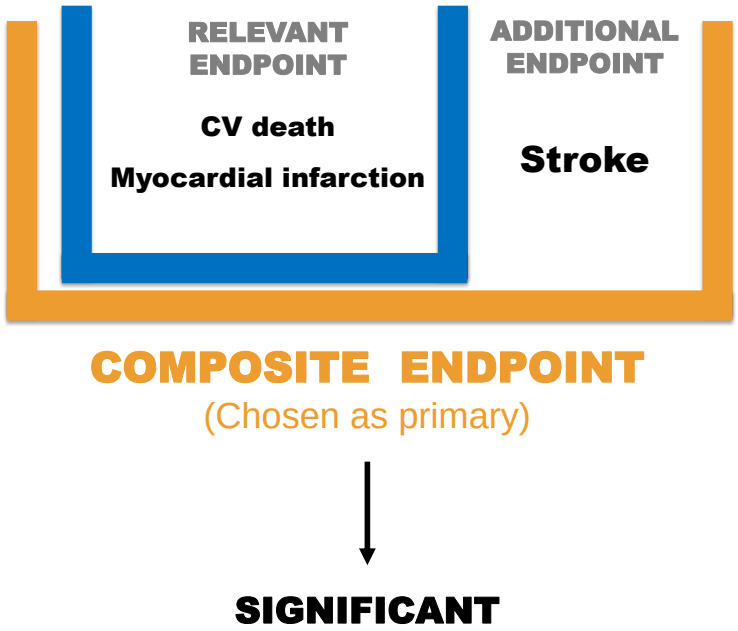
1. **CANCER CLINICAL TRIALS**
   DP: Disease progression
   OS: Overall survival
   PFS: Progression-free survival

2. **CARDIOVASCULAR DISEASE STUDIES**
   Cardiovascular death, myocardial infarction, stroke
   Hospitalization
   MACE: Major Adverse Cardiovascular Events

3. **HIV STUDIES**
   Virological failure
   Initiation of new treatment due to intolerance/toxicity
   TLOVR: Time to loss of virological response

Choosing the Primary Endpoint: An important decision

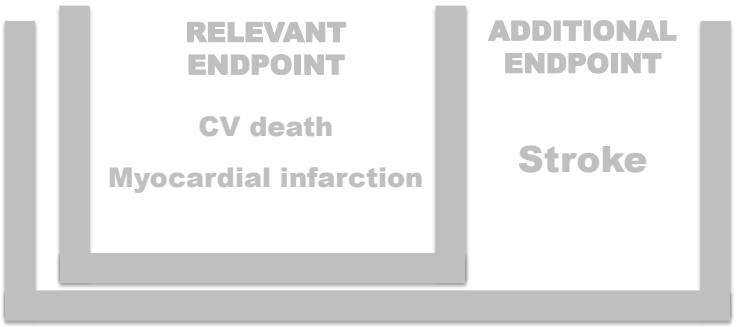
- **LIFE[1] study:**
  - Control group ($n = 4588$)
  - Losartan ($n = 4605$)



[1] Dahlöf B et al. Cardiovascular morbidity and mortality in the Losartan Intervention For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol (2002). *Lancet*, 359:995–1003.

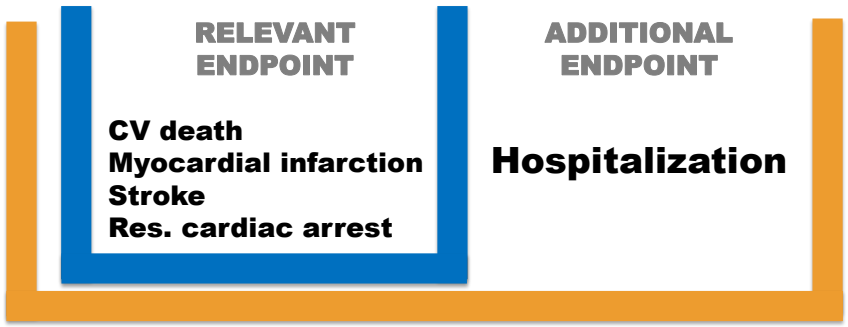
Choosing the Primary Endpoint: An important decision

- **LIFE[1] study:**
  - Control group ($n = 4588$)
  - Losartan ($n = 4605$)
- **ARISE[2] trial:**
  - Control group ($n = 3066$)
  - Succinobucol ($n = 3078$)



COMPOSITE ENDPOINT
(Chosen as primary)

SIGNIFICANT



COMPOSITE ENDPOINT
(Chosen as primary)

NON SIGNIFICANT

[1] Dahlöf B et al. Cardiovascular morbidity and mortality in the Losartan Intervention For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol (2002). *Lancet*, 359:995–1003.

[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

Goals of the Talk

1. **Statistical methodology (ARE) to guide the choice of the primary endpoint**
2. **CompARE: Web platform to facilitate the decision between CE and RE as the primary endpoint of the RCT**
3. **Extension to Binary Composite Endpoints**. Preliminar ideas

Goals of the Talk

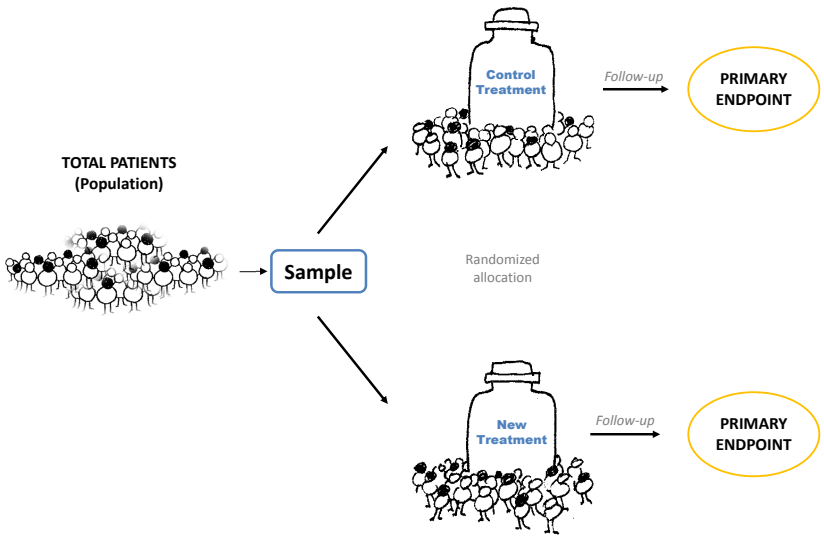
1. **Statistical methodology (ARE) to guide the choice of the primary endpoint**
2. **CompARE: Web platform to facilitate the decision between CE and RE as the primary endpoint of the RCT**
3. **Extension to Binary Composite Endpoints**. Preliminar ideas

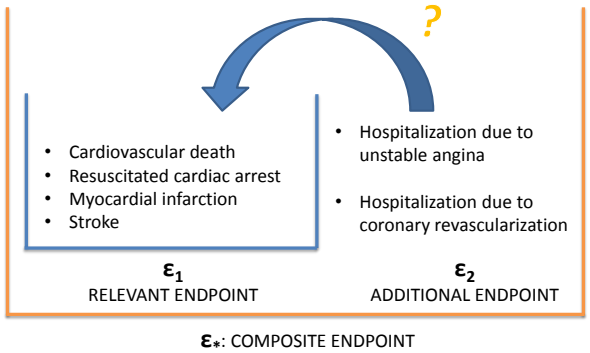
Goals of the Talk

1. **Statistical methodology (ARE) to guide the choice of the primary endpoint**
2. **CompARE: Web platform to facilitate the decision between CE and RE as the primary endpoint of the RCT**
3. **Extension to Binary Composite Endpoints**. Preliminar ideas

CONTEXT AND NOTATION

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





- 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.

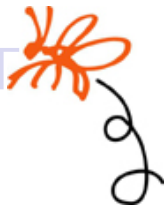
TESTING THE TREATMENT EFFECT: TWO SETS OF HYPOTHESIS

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



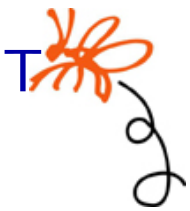
TESTING THE TREATMENT EFFECT: TWO SETS OF HYPOTHESIS

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



ARE methodology

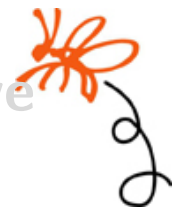
1. Testing H_0 vs H_1: Logrank test Z for T_1
   - Distinction of censoring cases
2. Testing H_0^* vs H_1^*: Logrank test Z_* for T_*
   - Copula Model for (T_1, T_2)
3. Asymptotic Relative Efficiency of Z_* versus Z: $\mathrm{ARE}(Z_*, Z)$ [(1)]
   - Representation of $\mathrm{ARE}(Z_*, Z)$ in terms of anticipatable parameters
   - ARE as ratio of sample sizes
   - Decision: robust with respect to the copula chosen
4. CompARE: Web Platform to facilitate computations

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

Censoring cases: is death one of the components?

- $\mathcal{E}_2$ does not contain death
  - 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_* censored by C
  - equal censoring in treatment groups
- $\mathcal{E}_2$ contains death
  - 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_* censored by C
  - Unequal censoring in treatment groups when treatment affects T_2

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



Censoring cases: is death one of the components?

- $\mathcal{E}_2$ does not contain death
  - 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_* censored by C
  - equal censoring in treatment groups
- $\mathcal{E}_2$ contains death
  - 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_* censored by C
  - Unequal censoring in treatment groups when treatment affects T_2

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

Z: Logrank test for T_1 (depends on censoring case)

- **Cases 1-3**: $\lambda_1^{(0)}(t)$, $\lambda_1^{(1)}(t)$ marginal hazards for T_1
- **Cases 2-4**: $\lambda_{C1}^{(0)}(t)$, $\lambda_{C1}^{(1)}(t)$ cause-specific hazards for T_1 when T_2 is a competing cause for T_1

- $H_0 : \mathrm{HR}(t) = \frac{\lambda_1^{(1)}(t)}{\lambda_1^{(0)}(t)} = 1 \Leftrightarrow$ NO EFFECT on T_1 (cases 1-3)
- Logrank $Z \sim N(0,1)$ under H_0

Z: Logrank test for T_1 (depends on censoring case)

- **Cases 1-3**: $\lambda_1^{(0)}(t)$, $\lambda_1^{(1)}(t)$ marginal hazards for T_1
- **Cases 2-4**: $\lambda_{C1}^{(0)}(t)$, $\lambda_{C1}^{(1)}(t)$ cause-specific hazards for T_1 when T_2 is a competing cause for T_1

- $H_0 : \mathrm{HR}(t) = \frac{\lambda_1^{(1)}(t)}{\lambda_1^{(0)}(t)} = 1 \Leftrightarrow$ NO EFFECT on T_1 (cases 1-3)
- Logrank $Z \sim N(0,1)$ under H_0

Z: Logrank test for T_1 under H_1

- View $\lambda_1^{(0)}(\cdot)$ as fixed, let $\lambda_{1,n}^{(1)}(\cdot)$ vary with n, and define the sequence:

  $$H_{1,n}: \log \mathrm{HR}_n(t) = \log\left(\frac{\lambda_{1,n}^{(1)}(t)}{\lambda_1^{(0)}(t)}\right) = \frac{g(t)}{\sqrt{n}}$$

- $Z \sim N(\mu, 1)$ [1] where

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

  - $p(t) = P_{H_0}(X = 1 | U \geq t)$
  - $V(t) = P_{H_0}(U \geq t)\lambda_1^{(0)}(t)dt = P_{H_0}(T_1 > t, C \geq t)\lambda_1^{(0)}(t)dt$ null sub-density function of observing a T_1 event at time t.

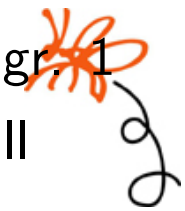
[1] Lagakos S.W. and Schoenfeld, D. Properties of Proportional-Hazards Score Tests under Misspecified Regression Models (1984). *Biometrics*, **40**, 1037–1048.

Z_*: Logrank test 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)

ARE TO ASSESS RELATIVE EFFICIENCY BETWEEN $\mathcal{E}_1$ VERSUS COMPOSITE $\mathcal{E}^* = \mathcal{E}_1 \cup \mathcal{E}_2$

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

We will assume:

- Equal number of subjects in the two treatment groups.
- End-of-study censoring C at time τ is the only noninformative censoring cause
- C identical across groups.
- HR_1 and HR_2: Constant treatment hazard ratios for T_1 and T_2.

ARE derived in terms of interpretable parameters

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

- **It 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 $\mathcal{E}_1$
- **It 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 $\mathcal{E}_2$

Interpretation of ARE. Criterion for Decision

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

$\boxed{\mathrm{ARE} \approx \frac{n}{n_*}} \Rightarrow$ Usual interpretation of ARE holds:

Given $0 < \alpha < \Pi < 1$,

$$\lim_{\substack{\mathrm{HR}_{1,n}(t)\to 1 \\ \mathrm{HR}_{2,n}(t)\to 1}} \frac{n}{n_*} = \mathrm{ARE}(Z_*, Z).$$

where n and n_*: sample sizes required for Z_n and $Z_{n_*}^*$ to have power $\geq \Pi$ at level α.

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

Summary of method

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 $p_1, p_2, \mathrm{HR}_1, \mathrm{HR}_2, \rho$

Copula model for (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

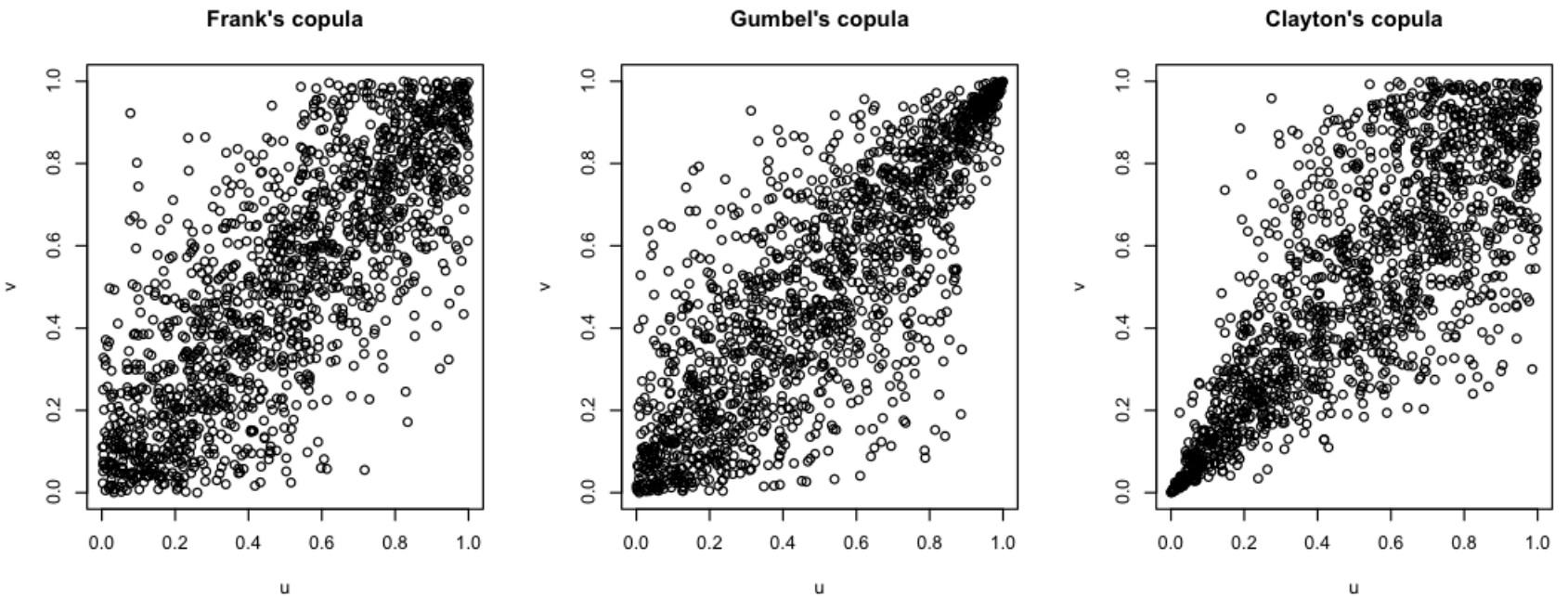


Figure: Common bivariate copulas

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 $T_* = \min\{T_1, T_2\}$

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

ARE Comparison for Frank, Gumbel and Clayton copulas

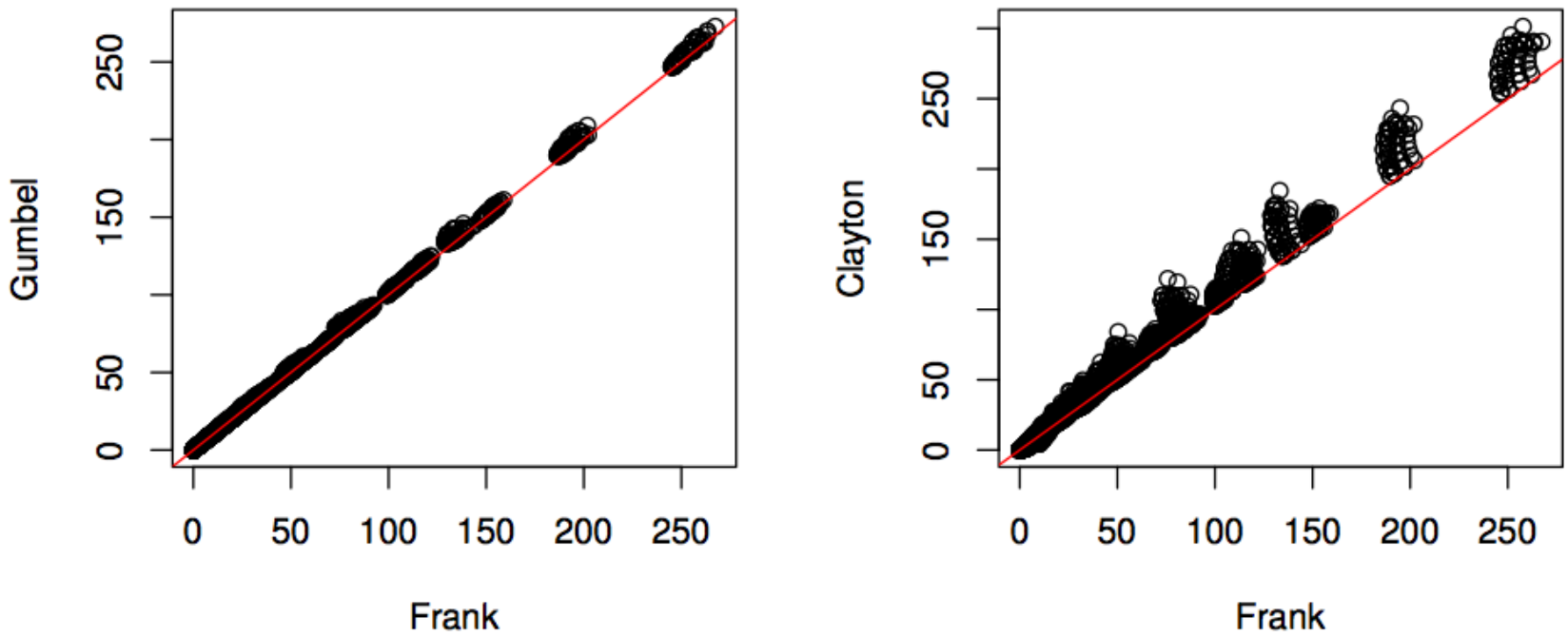


Figure: Pairwise ARE correlations based on 72576 simulated situations

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

Plana, O. and Gómez G. Selecting the primary endpoint in a randomized clinical trial. The ARE method. (Submitted)

Robustness w.r.t. choice of the copula

	$ARE_{Gumbel} > 1$	$ARE_{Gumbel} \leq 1$	$ARE_{Clayton} > 1$	$ARE_{Clayton} \leq 1$
$ARE_{Frank} > 1$	59.5%	0.02%	59.2%	0.4%
$ARE_{Frank} \leq 1$	1.9%	38.5%	4.9%	35.6%

Degree of agreement **Frank - Gumbel** $\rightarrow$ 98.0%
Degree of agreement **Frank - Clayton** $\rightarrow$ 94.7%

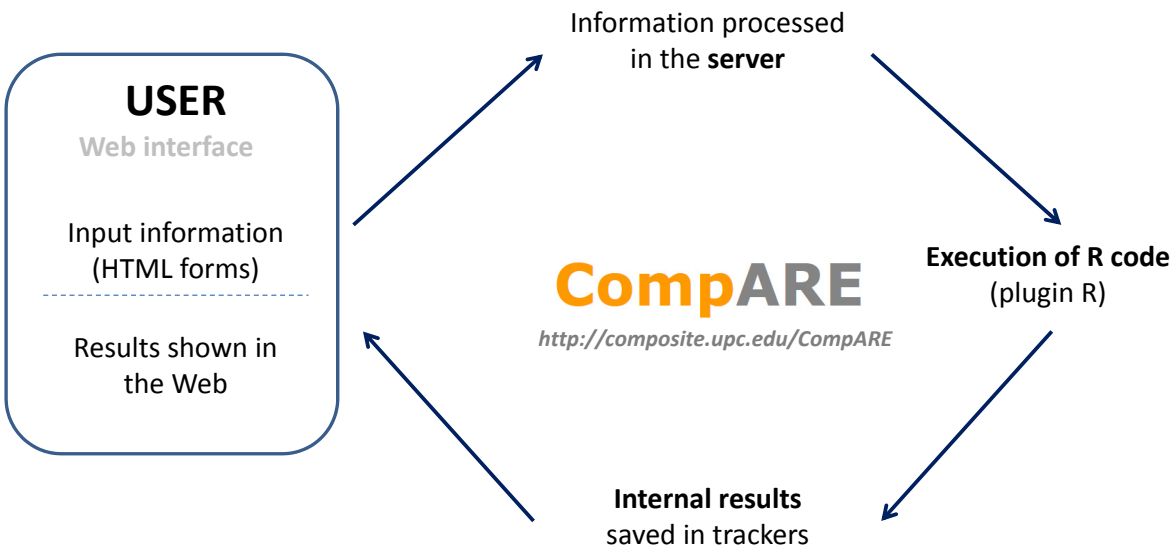
DISCORDANT CASES =7%

Discordant cases	**n**	**mean (SD)**	**min**	Q_1	**median**	Q_3	P_{95}	**max**
$\lvert ARE_F - ARE_G \rvert$	1426	0.04 (0.03)	0.004	0.02	0.05	0.06	0.11	0.14
$\lvert ARE_F - ARE_C \rvert$	3812	0.11 (0.08)	0.001	0.04	0.09	0.17	0.27	0.36

ONLY 1.6% cases with $|ARE_F - ARE_C| > 0.15$ corresponding to $HR_1 = HR_2$ or $HR_1 = HR_2 - 0.1$ and $\rho \geq 0.45$

CompARE interface

- 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



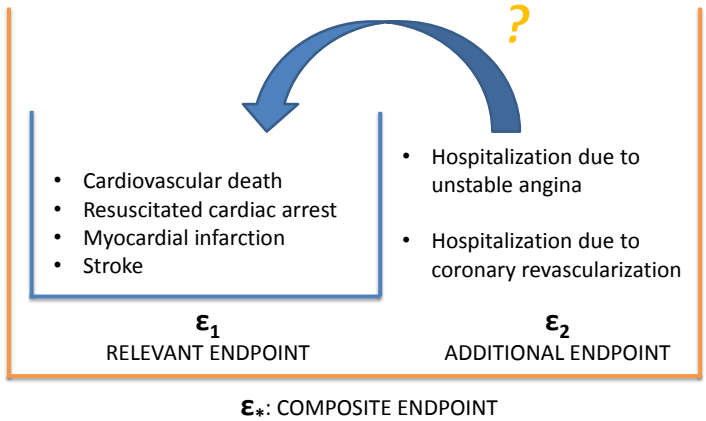
Software used to built the Interface

- **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 after an acute coronary syndrome with succinobucol (Tardif *et al.* Lancet 2008)

6144 patients randomized to receive succinobucol in addition to SOC:

0. SOC ($n = 3066$)
   - 252 events of $\mathcal{E}_1$
   - 277 events of $\mathcal{E}_2$
   - 529 events of $\mathcal{E}_*$
1. succinobucol ($n = 3078$)
   - 207 events of $\mathcal{E}_1$
   - 323 events of $\mathcal{E}_2$
   - 530 events of $\mathcal{E}_*$



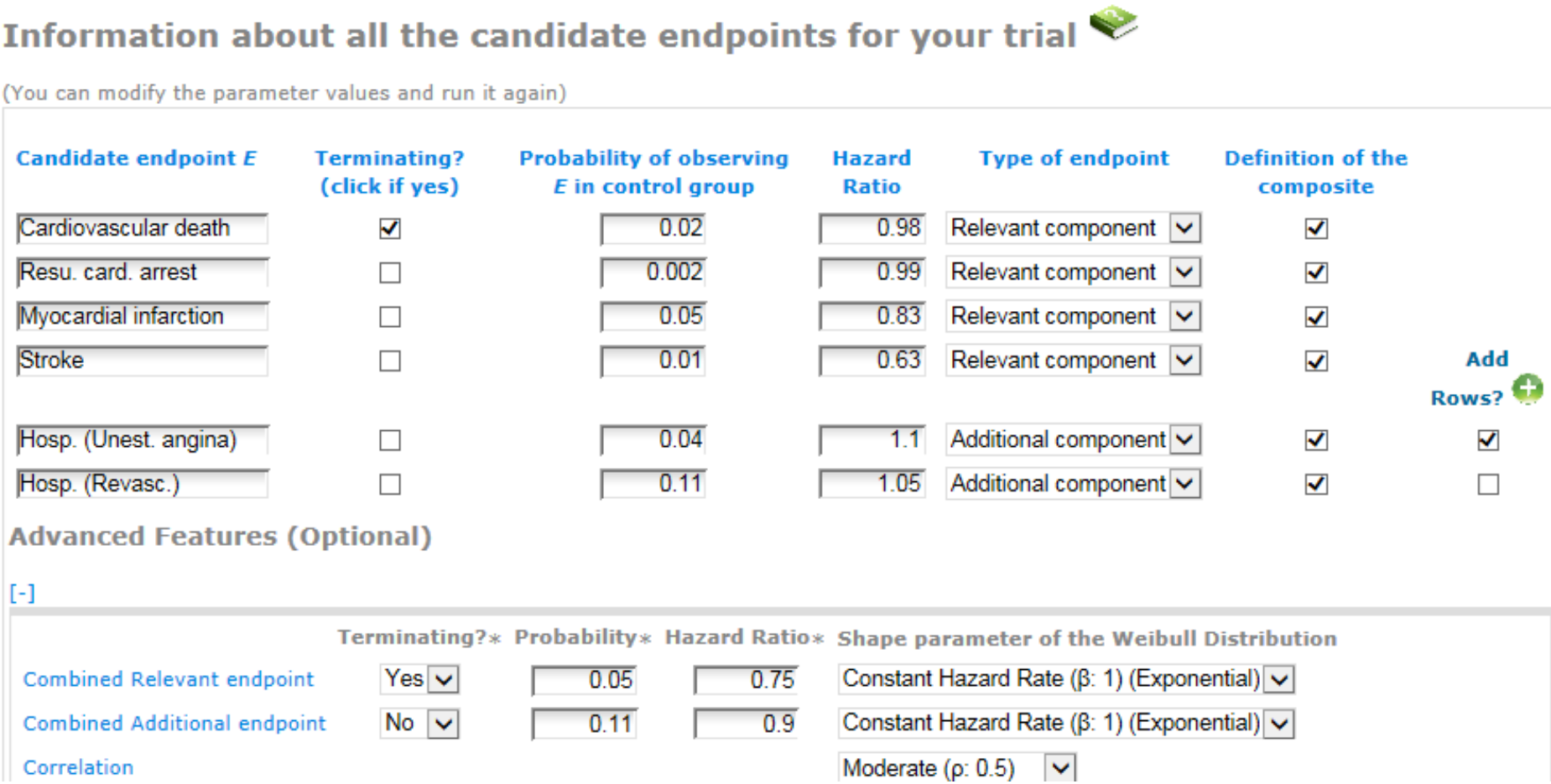
- $\text{Prob}\{\mathcal{E}_1\} = 0.08$, $\text{HR}_1 = 0.81$
- $\text{Prob}\{\mathcal{E}_2\} = 0.09$, $\text{HR}_2 = 1.05$

- Beneficial effect of succinobucol ($p = 0.029$) on $\mathcal{E}_1$
- **Failed to show** significant differences on $\mathcal{E}_*$.
- Hospital admission component MASKED the mortality effect

Analysis with CompARE

Relevant endpoints: CV death, Resusc CA, MI, Stroke

Additional endpoints: Hospitalizations



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	Add Rows?
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	☑	
Hosp. (Unest. angina)	☐	0.04	1.1	Additional component	☑	☑
Hosp. (Revasc.)	☐	0.11	1.05	Additional component	☑	☐

Advanced Features (Optional)

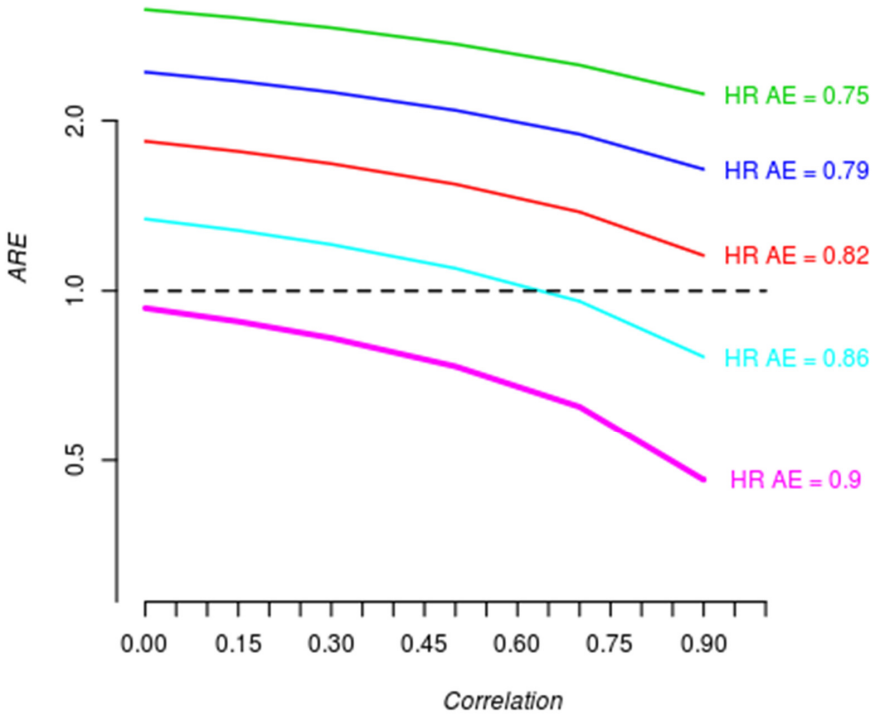
[-]

	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)

Analysis with CompARE

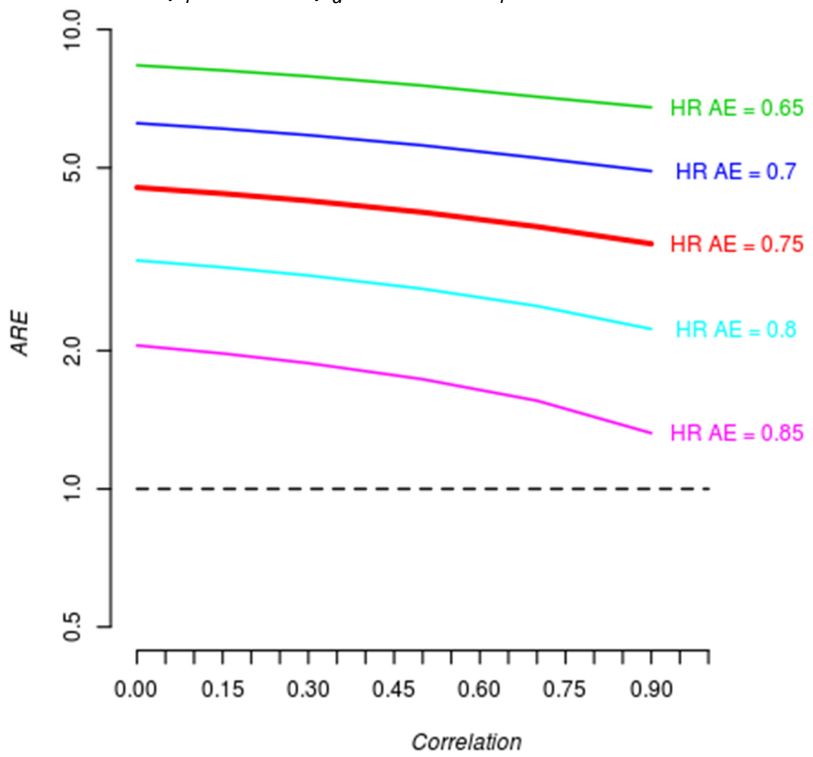
Fixed parameters:

$p_r = 0.05$ $p_a = 0.11$ $HR_r = 0.75$

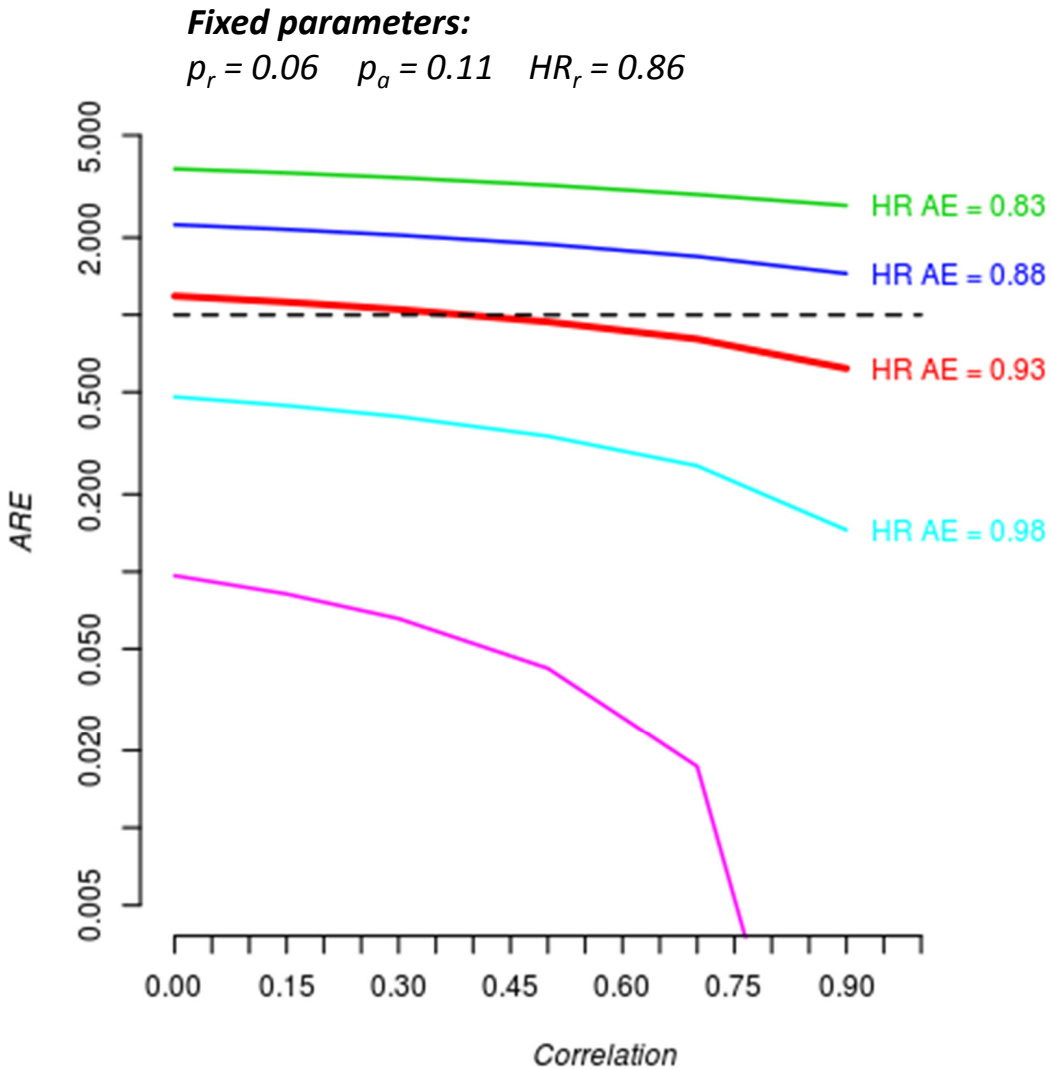


Fixed parameters:

$p_r = 0.05$ $p_a = 0.11$ $HR_r = 0.8$



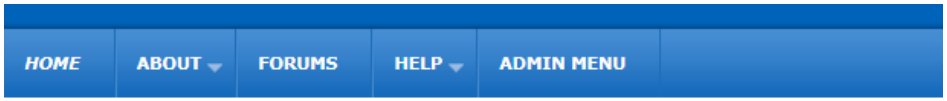
$\mathcal{E}_*$ would have been justified if $\mathrm{HR}_2 \leq 0.88$

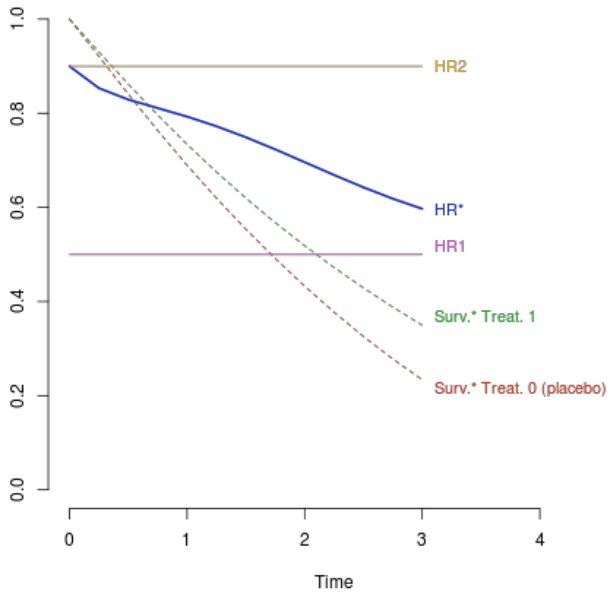


$\mathrm{HR}_2 = 1.05 \Rightarrow \mathrm{ARE}(T_* \mathrm{vs} T_1) < 1, \ \forall\ \rho(T_1, T_2) \Rightarrow \mathcal{E}_1$ should have been used.

Other outputs

- **Survival and Hazard Ratio functions**





- **Numerical results in 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

- **Reported recommendations in text**
- **List of previous results**

Binary Composite Endpoints. Instances in HIV

TEMPTATIVE PRIMARY BINARY ENDPOINTS

1.
   - Relevant: Virologic failure (increase in plasma HIV of RNA greater than 200 copies/ml)
   - Additional: Lost to Follow Up/ Initiation of new treatment due to toxicity /Intolerance/ Death
   - LOVR (Loss of Virological Response): Virologic failure OR Lost to Follow Up/ Initiation of new treatment due to toxicity /Intolerance/ Death

2.
   - RE: Virologic failure (efficacy)
   - AE: Adverse effects (safety)
   - Binary CE: Virologic failure OR Adverse effects

3.
   - RE: CD4 cell < 250
   - AE: Initiation of Antiretroviral therapy
   - Binary CE: CD4 cell < 250 OR Initiation of Antiretroviral therapy

Binary Composite Endpoints. Instances in HIV

TEMPTATIVE PRIMARY BINARY ENDPOINTS

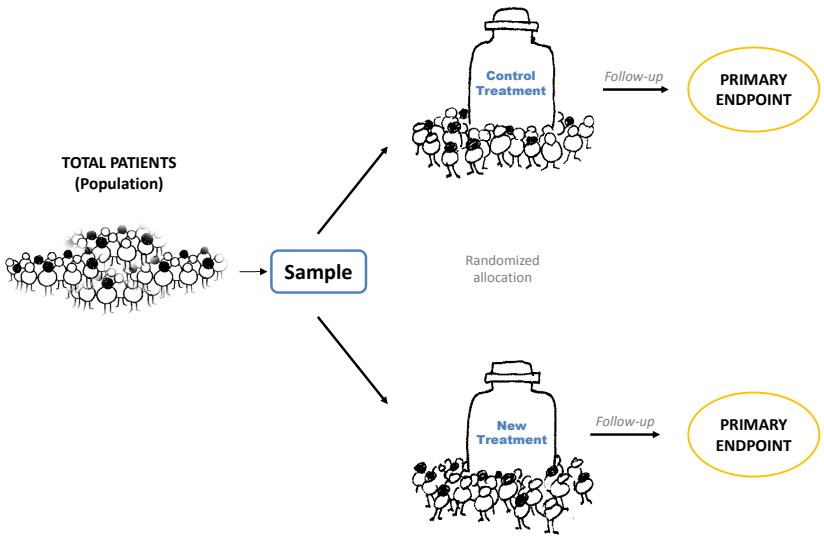
1.
   - Relevant: Virologic failure (increase in plasma HIV of RNA greater than 200 copies/ml)
   - Additional: Lost to Follow Up/ Initiation of new treatment due to toxicity /Intolerance/ Death
   - LOVR (Loss of Virological Response): Virologic failure OR Lost to Follow Up/ Initiation of new treatment due to toxicity /Intolerance/ Death

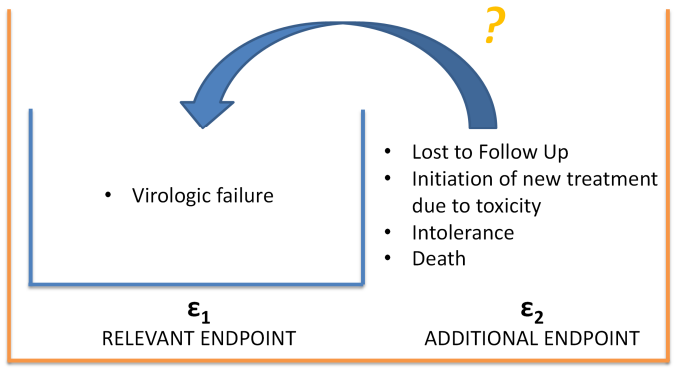
2.
   - RE: Virologic failure (efficacy)
   - AE: Adverse effects (safety)
   - Binary CE: Virologic failure OR Adverse effects

3.
   - RE: CD4 cell < 250
   - AE: Initiation of Antiretroviral therapy
   - Binary CE: CD4 cell < 250 OR Initiation of Antiretroviral therapy

CONTEXT AND NOTATION

RCT for comparing the efficacy of treatment 1 versus treatment 0





ε_*: COMPOSITE ENDPOINT
LOVR (Loss of Virological Response)

- RELEVANT ENDPOINT $\mathcal{E}_1$:
  $Y_{ij1} = \mathbf{1}\{RE\}$ (indiv j, group i)
- ADDITIONAL ENDPOINT $\mathcal{E}_2$:
  $Y_{ij2} = \mathbf{1}\{AE\}$ (indiv j, group i)
- COMPOSITE ENDPOINT $\mathcal{E}_1 \cup \mathcal{E}_2$:

$$Y_{ij*} = \begin{cases} 1 & \text{if } Y_{ij1} + Y_{ij2} \geq 1 \\ 0 & \text{if } Y_{ij1} + Y_{ij2} = 0 \end{cases}$$

Notation

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

Relationship between p_{i1}, p_{i2} and p_{i*} via Bahadur's general form

Bahadur's theorem

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 $q_{ik} = 1 - p_{ik}$.

Definition of p_{i*}

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

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

Hypothesis of no treatment effect

Null Hypothesis

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

Equivalent null hypothesis???

$H_0 : p_{01} = p_{11} \not\Leftrightarrow H_0^* : p_{0*} = p_{1*}$

However, if $p_{01} = p_{11}$ and $p_{02} = p_{12}$ and $\rho_0 = \rho_1 \implies H_0^* : p_{0*} = p_{1*}$

Assumption: $\rho_0 = \rho_1 = \rho \rightsquigarrow$ Reasonable

Two Sample Binomial test statistics

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

- $\tilde{p}_1 = \frac{Y_{01}+Y_{11}}{N_0+N_1}$, common consistent estimator of p_{01} and 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 $H_0^* : p_{0*} = p_{1*}$

- $\tilde{p}_* = \frac{Y_{0*}+Y_{1*}}{N_0+N_1}$, common consistent estimator of p_{0*} and 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^*

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

Two Sample Binomial test statistics

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

- $\tilde{p}_1 = \frac{Y_{01}+Y_{11}}{N_0+N_1}$, common consistent estimator of p_{01} and 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 $H_0^* : p_{0*} = p_{1*}$

- $\tilde{p}_* = \frac{Y_{0*}+Y_{1*}}{N_0+N_1}$, common consistent estimator of p_{0*} and 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^*

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

Asymptotic relative efficiency of T_* versus T_1

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

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

where $\mathrm{O}_{01} = p_{01}/1 - p_{01}$, $\mathrm{O}_{02} = p_{02}/1 - p_{02}$.

The ARE as a Function of Interpretable Parameters

- p_{01} and $p_{02} \rightarrow$ Probability exhibiting the RE and AE in control group.
- OR_1 and OR_2
- $\rho \rightarrow$ The correlation between RE and AE

SUMMARIZING

- ARE: Conceptual framework as a tool to decide whether or not a CE should be used when comparing two treatment groups in a RCT
- Use of Composite Endpoints has to be justified from a clinical point of view
- Careful study of the anticipated values for (p_1, p_2, HR_1) and the corresponding ARE in the planning proces of any RCT
- **CompARE** to compute the ARE for time-to-event endpoints
- Extending **CompARE** to sample size computation.
- ARE for binary CE.
- Extending **CompARE** to binary CE.

Thanks to my coauthors



Oleguer, Susana and Nuria in EMR-IBS 2013, Tel-Aviv


Thank you!


Thanks!


Thank
you!


THANK
YOU!


THANK
YOU

Some References

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

✓ Gómez G, Gómez-Mateu M, Dafni U.(2014). Informed Choice of Composite End Points in Cardiovascular Trials. *Circulation. Cardiovascular Quality and Outcomes*, **7**, 170–178.

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

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

✓ Plana, O. and Gómez G. (2014) Selecting the primary endpoint in a randomized clinical trial. The ARE method. (Submitted)

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