Sample size recalculation for a skewed outcome in two-stage three-arm sequential non-inferiority clinical trials: a simulation study

Maria Vittoria Chiaruttini, Ph.D Candidate in Biostatistics

University of Padua, Italy

Workshop on Adaptive Designs and Multiple Testing Procedures

19-21 April 2023 - Novartis Campus, Basel

Università degli Studi di Padova

Dipartimento di Scienze Cardio-Toraco-Vascolari e Sanità Pubblica

Unità di Biostatistica, Epidemiologia e Sanità Pubblica

Background

- Non-inferiority clinical trial & FDA guidelines
- The base-point of this study: the Lanyu Lei resampling (original) algorithm
- The skewed outcomes in clinical research

Non-inferiority clinical trial & FDA guidelines (1)

ACTIVE TREATMENT →

Fewer side effects →

Less time to treat →

Less expensive →

NEW TREATMENT

Non-inferiority clinical trial & FDA guidelines (2)

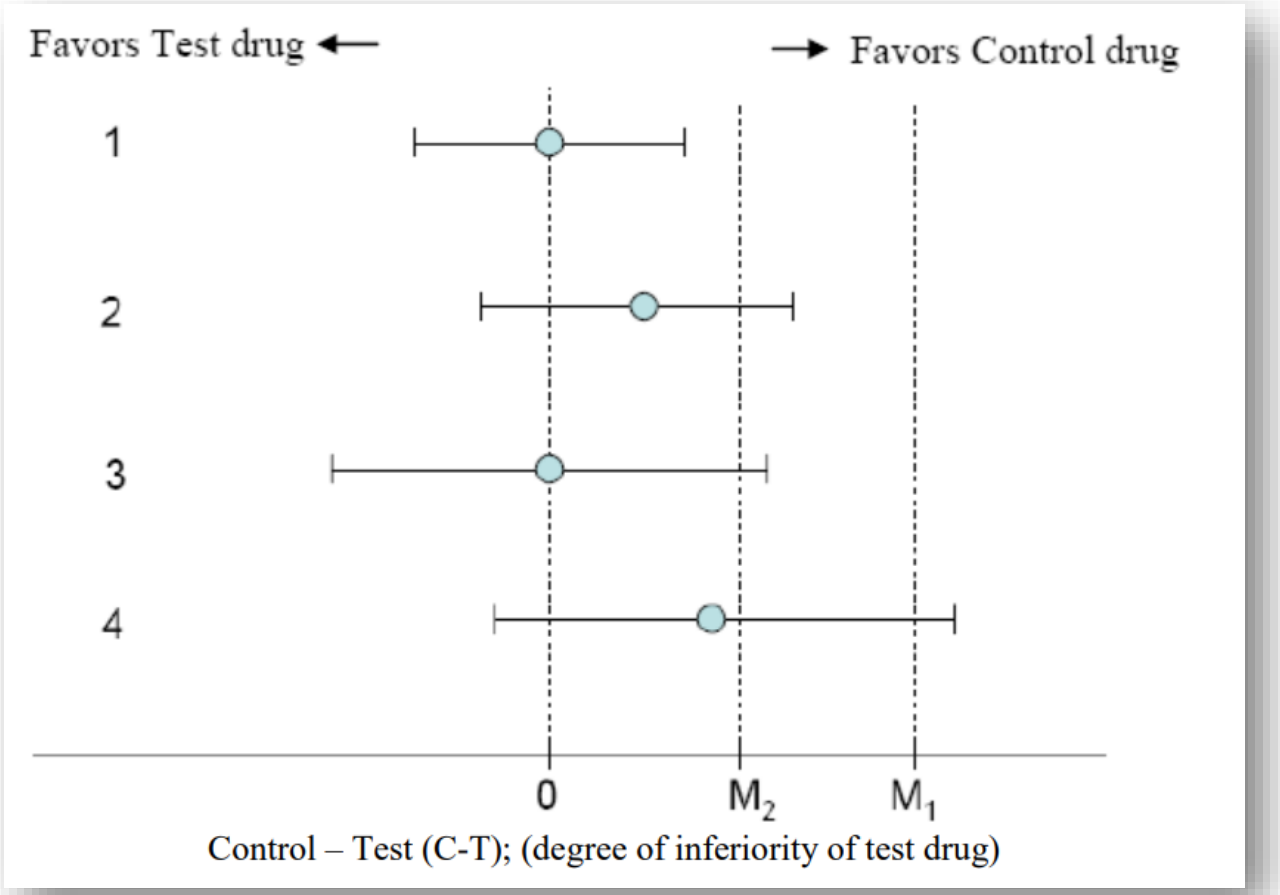
Non-Inferiority Clinical Trials to Establish Effectiveness

Guidance for Industry

Additional copies are available from:
Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
and/or
Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010
Email: ocod@fda.hhs.gov
http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

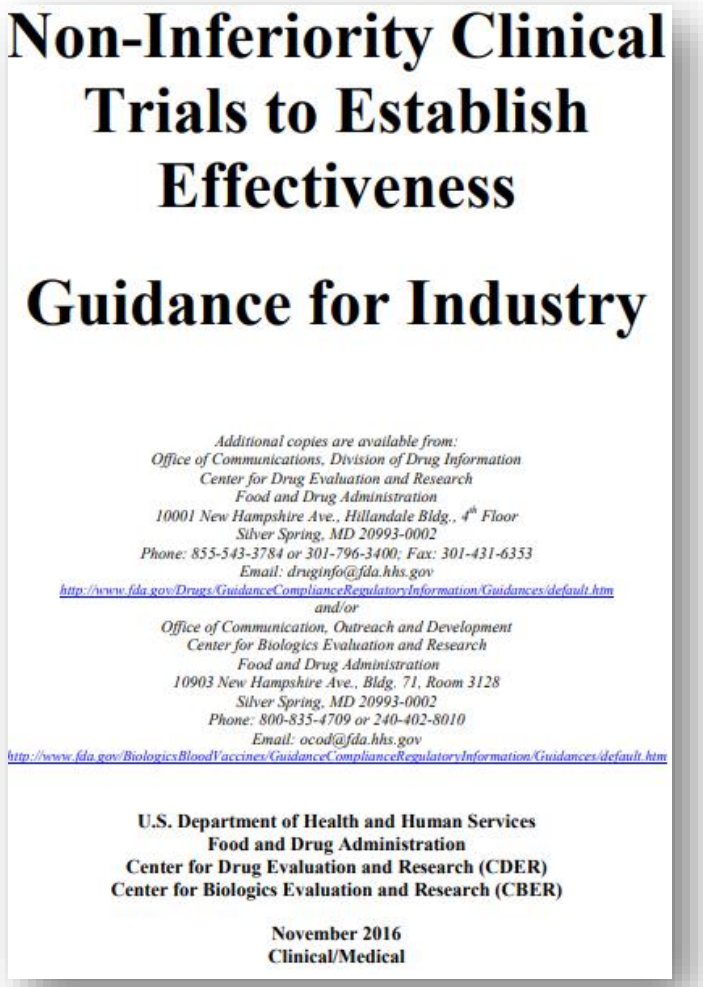
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

November 2016
Clinical/Medical



Control – Test (C-T); (degree of inferiority of test drug)

Non-inferiority clinical trial & FDA guidelines (3)



Non-Inferiority Clinical Trials to Establish Effectiveness

Guidance for Industry

Additional copies are available from:
Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
and/or
Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010
Email: ocod@fda.hhs.gov
http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

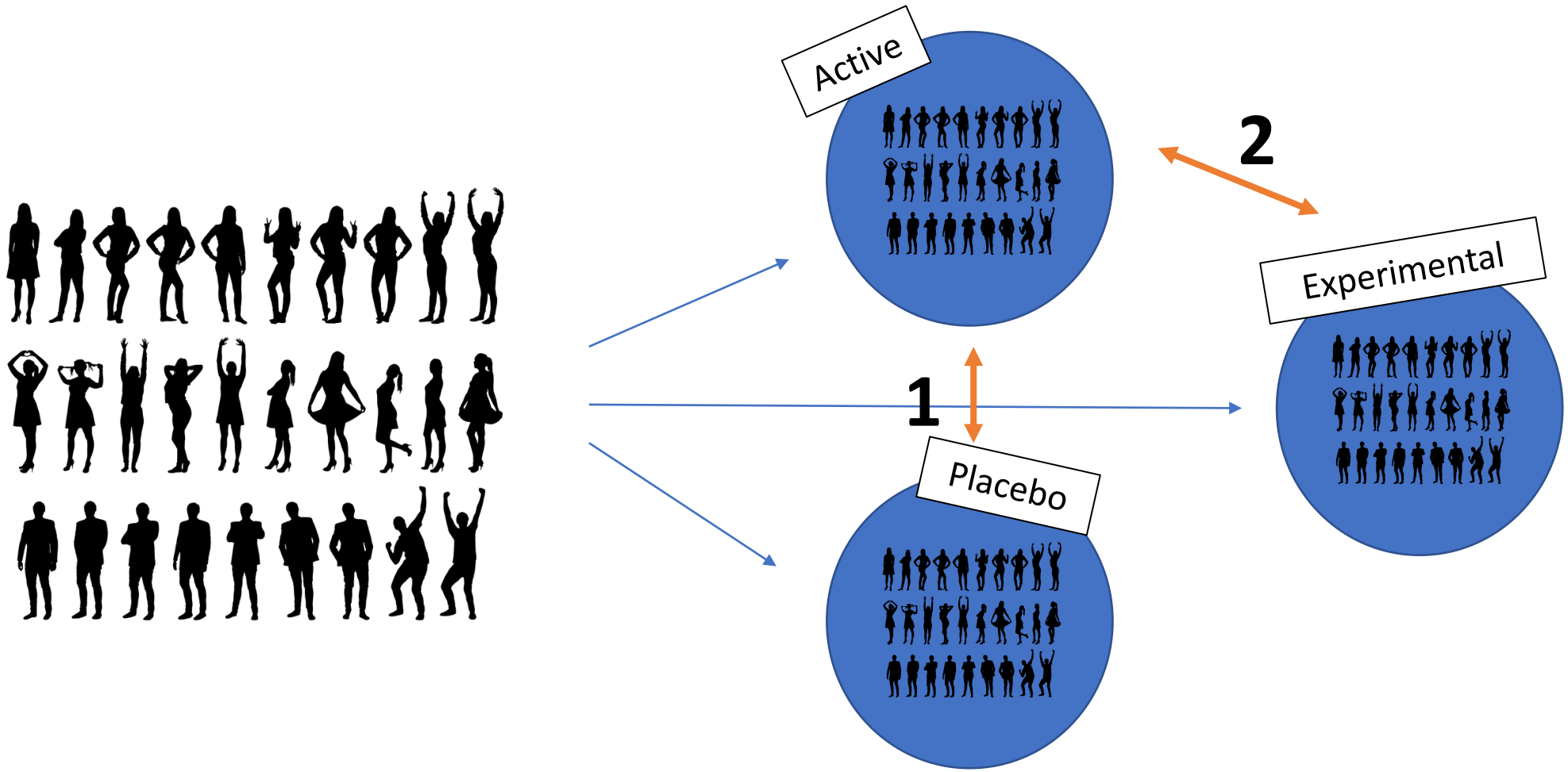
November 2016
Clinical/Medical

- **Constancy** → The demonstrated effect of the active control over placebo in the historical trial has not changed over time

- **Assay sensitivity** → The study's ability to distinguish an effective treatment from a less effective or an ineffective one

This assumptions are not testable in a trial without a concurrent placebo group

Non-inferiority clinical trial & FDA guidelines (4)



Non-inferiority clinical trial & FDA guidelines (5)

Superiority

$$H_{0,AP}^{(s)}: \mu_A \leq \mu_P \text{ versus } H_{1,AP}^{(s)}: \mu_A > \mu_P$$

Noninferiority

Fixed Margin

or

$$H_{0,TA}^{(n)}: \mu_T \leq \mu_A - \Delta \text{ versus } H_{1,TA}^{(n)}: \mu_T > \mu_A - \Delta$$

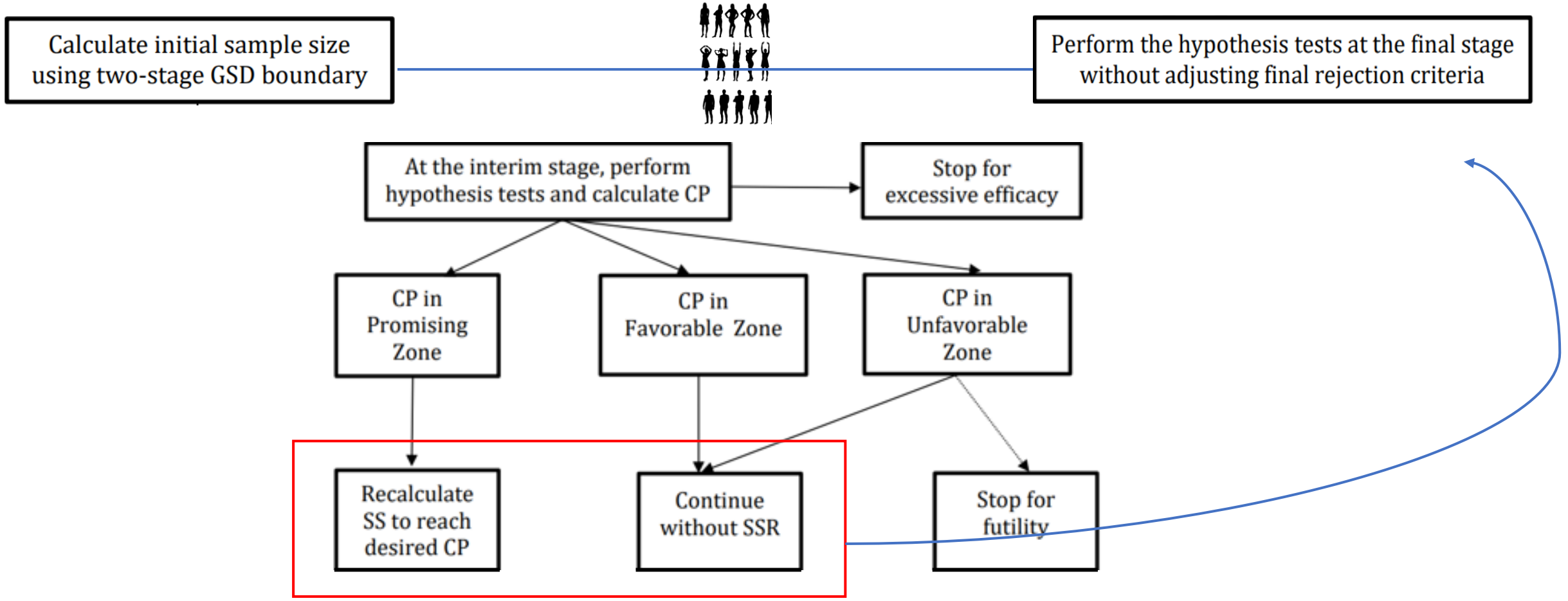
Noninferiority

Effect of preservation

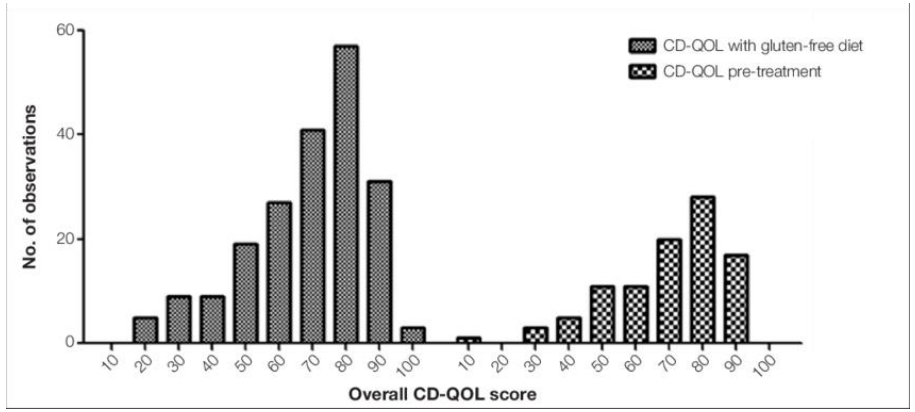
$$H_{0,TA}^{(n)}: \frac{\mu_T - \mu_P}{\mu_A - \mu_P} \leq 1 - \lambda \text{ versus } H_{1,TA}^{(n)}: \frac{\mu_T - \mu_P}{\mu_A - \mu_P} > 1 - \lambda$$

The base-point of this study: the Lanyu Lei resampling algorithm

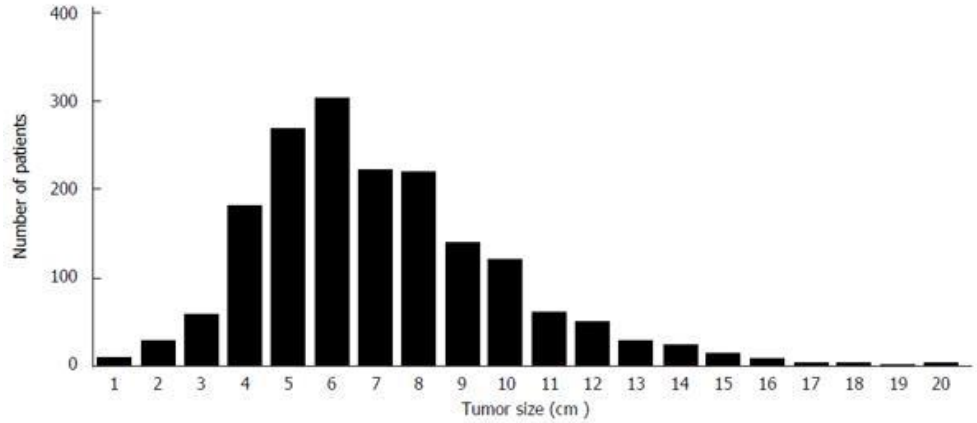
A group sequential design based on the promising zone of conditional power (1)



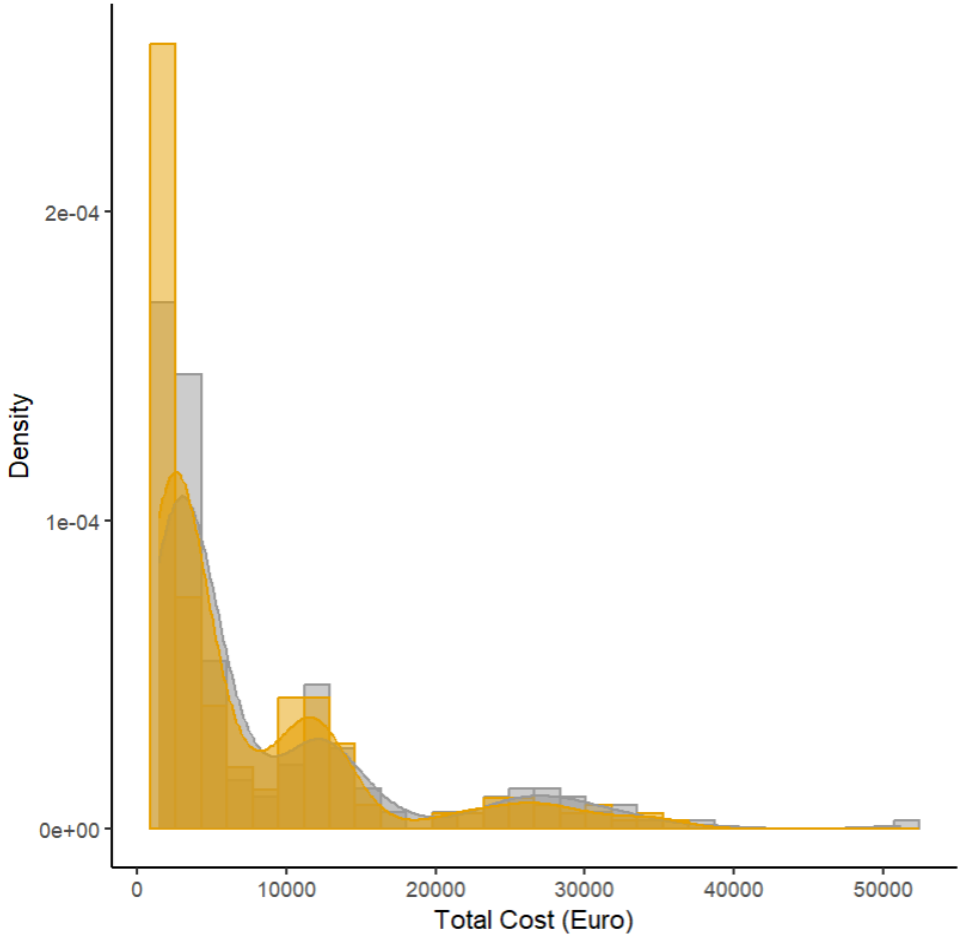
The skewed outcomes in clinical research



Casellas F. et al. (2014)



Huang B. et al. (2016)



Desideri A. et al. (2003)

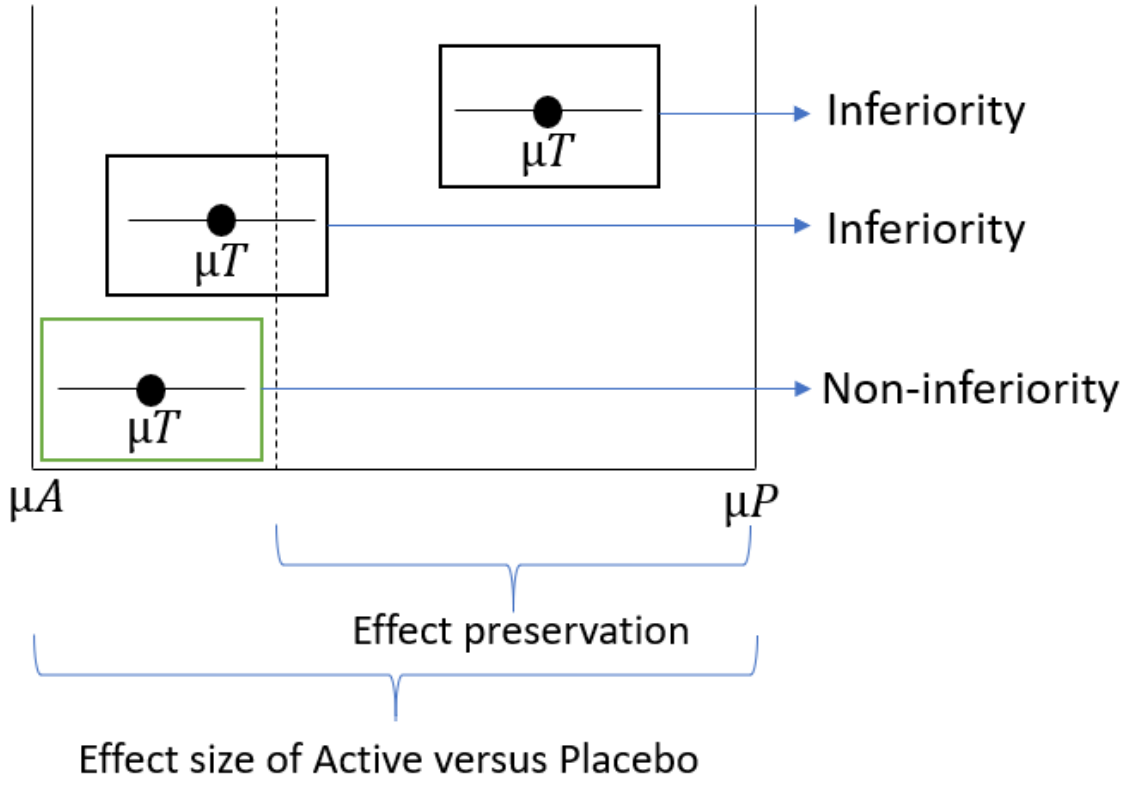
Methods

- The Coverage probability of the original sample size re-calculation method in the presence of deviation from normality
- The proposed algorithm for skewed outcomes
- A real-data example
- A brief presentation of the r-shiny web application

The coverage probability of the original sample size re-calculation method in the presence of deviation from normality

COVERAGE PROBABILITY:

The **total number** of **rejected nulls** over the **total number** of **simulations** (N=10,000), **under** the **alternative hypothesis**, that in case of "smaller is better" is:



The coverage probability of the original sample size re-calculation method in the presence of deviation from normality

Sample size

Simulated populations

$Y(P) \sim N(\mu P, \sigma)$

$Y(A) \sim N(\mu A, \sigma)$

$Y(T) \sim N(\mu T, \sigma)$

$Y(P) \sim Gamma(kP, \theta P)$

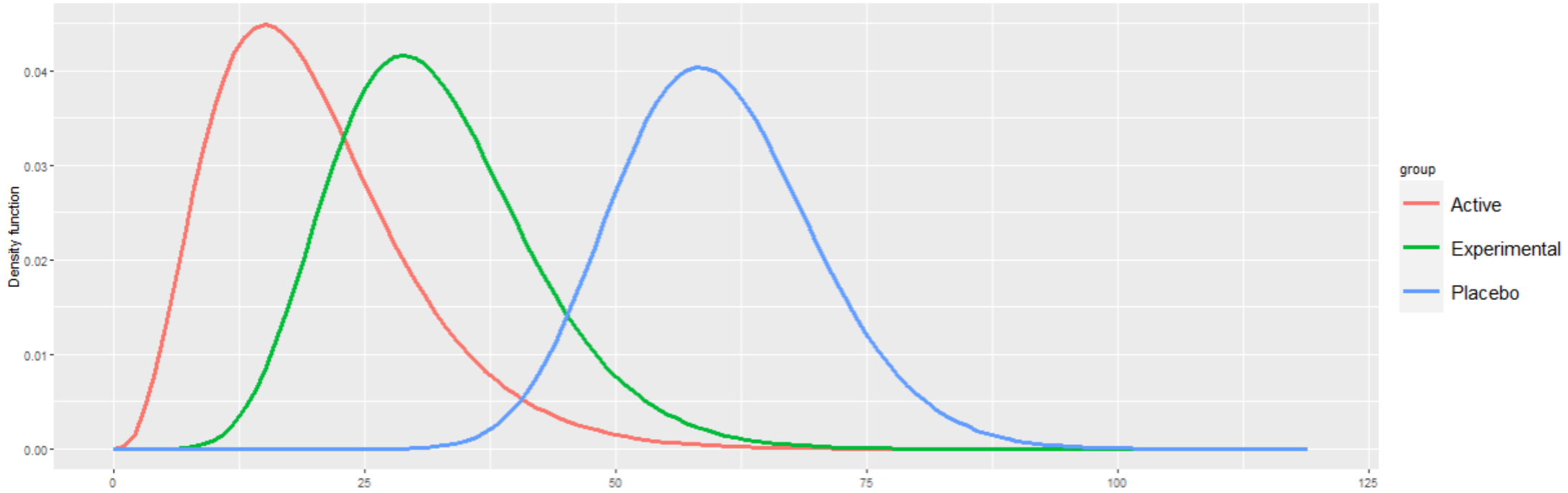
$Y(A) \sim Gamma(kA, \theta A)$

$Y(T) \sim Gamma(kT, \theta T)$

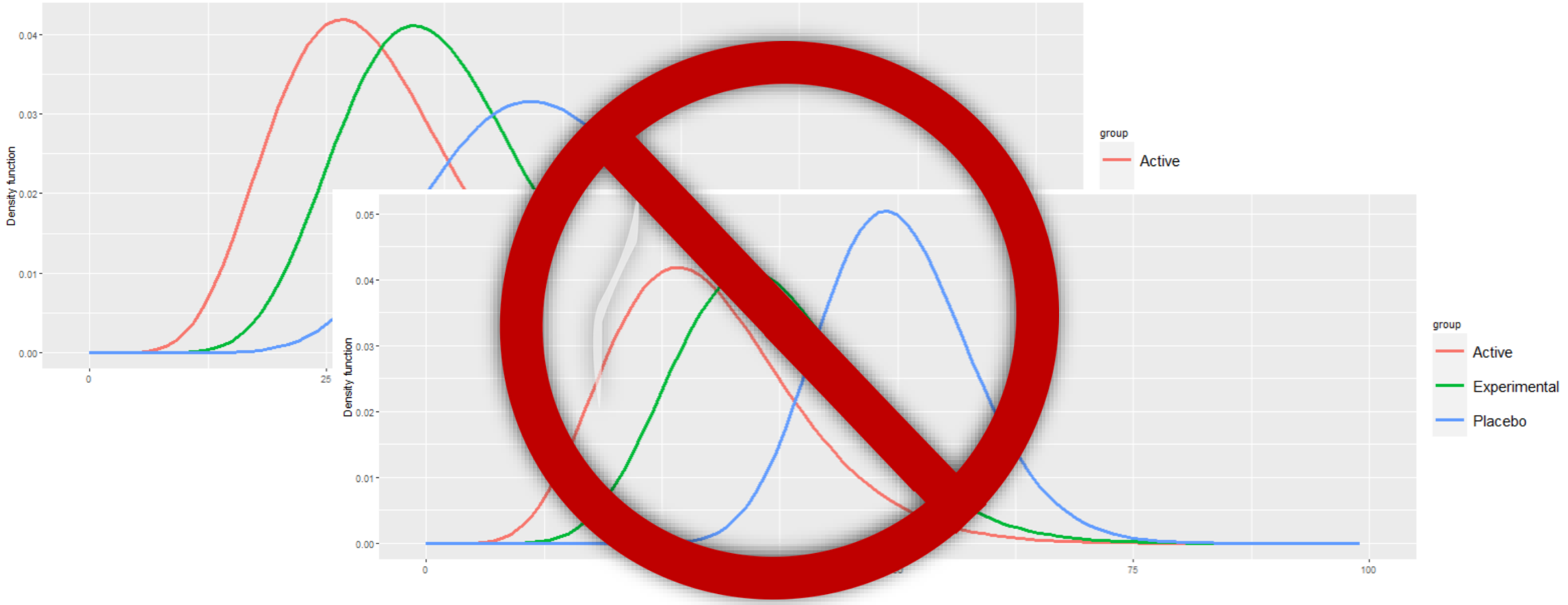
$$\begin{cases} \mu = k * \theta \\ \sigma = \sqrt{k * \theta^2} \end{cases} =$$

$$\begin{cases} k = (\mu/\sigma)^2 \\ \theta = \sigma^2/\mu \end{cases}$$

The coverage probability of the original sample size re-calculation method in the presence of deviation from normality



The coverage probability of the original sample size re-calculation method in the presence of deviation from normality



The proposed algorithm for skewed outcomes (1)

The power function used to estimate the initial sample size

$$1-\beta=\phi\left(\frac{\mu T-(1-\lambda)\mu A-\lambda\mu P)}{\frac{\sigma}{\sqrt{nT}}\sqrt{1+\frac{(1-\lambda)^2}{cA}+\frac{\lambda^2}{cP}}}-Z_{1-\alpha(2)}\right)$$

$$1-\beta=\phi\left(\frac{\mu T-(1-\lambda)\mu A-\lambda\mu P)}{\sqrt{\frac{Var(T)}{nT}+\frac{(1-\lambda)^2*Var(A)}{nT*cA}+\frac{\lambda^2 Var(P)}{nT*cP}}}-t_{1-\alpha(2),\nu}\right)$$

The proposed algorithm for skewed outcomes (2)

The test statistics for $H_{1,\,AP} = \mu A < \mu P$:

$$Z_{PA} = \frac{\mu P - \mu A}{\sigma\sqrt{\frac{1}{nT * cA} + \frac{1}{nT * cP}}}$$

$$T_{PA} = \frac{\mu P - \mu A}{\sqrt{\frac{Var(A)}{nT * cA} + \frac{Var(P)}{nT * cP}}}$$

The proposed algorithm for skewed outcomes (3)

- The test statistics for $H_{1,\,TA} = \frac{(\mu P - \mu T)}{(\mu P - \mu A)} > 1 - \lambda$:

$$Z_{TA} = \frac{(1-\lambda)\mu A + \lambda \mu P - \mu T}{\frac{\sigma}{\sqrt{nT}}\sqrt{1 + \frac{(1-\lambda)^2}{cA} + \frac{\lambda^2}{cP}}}$$

$$T_{TA} = \frac{(1-\lambda)\mu A + \lambda \mu P - \mu T}{\sqrt{\frac{Var(T)}{nT} + \frac{(1-\lambda)^2 * Var(A)}{nT * cA} + \frac{\lambda^2 Var(P)}{nT * cP}}}$$

The proposed algorithm for skewed outcomes (4)

The conditional power function:

$$CondP = 1 - \phi\left(\frac{Z_{1-\alpha(2)}\sqrt{nT(2)} - Z_{TA}(1)\sqrt{nT(1)}}{\sqrt{nT(2) - nT(1)}} - \frac{(1-\lambda)\mu A + \lambda\mu P - \mu T}{\frac{\sigma}{\sqrt{nT(2) - nT(1)}}\sqrt{1 + \frac{(1-\lambda)^2}{cA} + \frac{\lambda^2}{cP}}}\right)$$

$$CondP = 1 - \phi\left(\frac{t_{1-\alpha(2),\nu}\sqrt{nT(2)} - T_{TA}(1)\sqrt{nT(1)}}{\sqrt{nT(2) - nT(1)}} - \frac{(1-\lambda)\mu A + \lambda\mu P - \mu T}{\sqrt{\frac{Var(T)}{nT(2) - nT(1)} + \frac{(1-\lambda)^2 * Var(A)}{(nT(2) - nT(1)) * cA} + \frac{\lambda^2 Var(P)}{(nT(2) - nT(1)) * cP}}}\right)$$

A real data example (1)

European Heart Journal (2003) 24, 1630–1639

ELSEVIER

EUROPEAN SOCIETY OF CARDIOLOGY

Cost of strategies after myocardial infarction (COSTAMI)

A multicentre, international, randomized trial for cost-effective discharge after uncomplicated myocardial infarction

Alessandro Desideri[a*], Paolo Maria Fioretti[b], Lauro Cortigiani[c], Dario Gregori[d], Claudio Coletta[e], Carlo Vigna[f], Francesco Tota[g], Riccardo Rambaldi[h], Jeroen Bax[i], Leopoldo Celegon[j], Riccardo Bigi[k], Eugenio Picano[l], on behalf of the COSTAMI (Cost of Strategies After Myocardial Infarction) Trial Investigators

A real data example (2)

Groups	Total cost	
	Mean	**Standard Deviation**
Usual Care Strategy= Placebo/control group	8344	8894
Early Discharge=active comparator group	7391	7903

$\mu P - \mu A$ = € 953

$1 - \lambda = 0.5$

$\mu P - \mu T >$ 953*0.5 = € 476.5

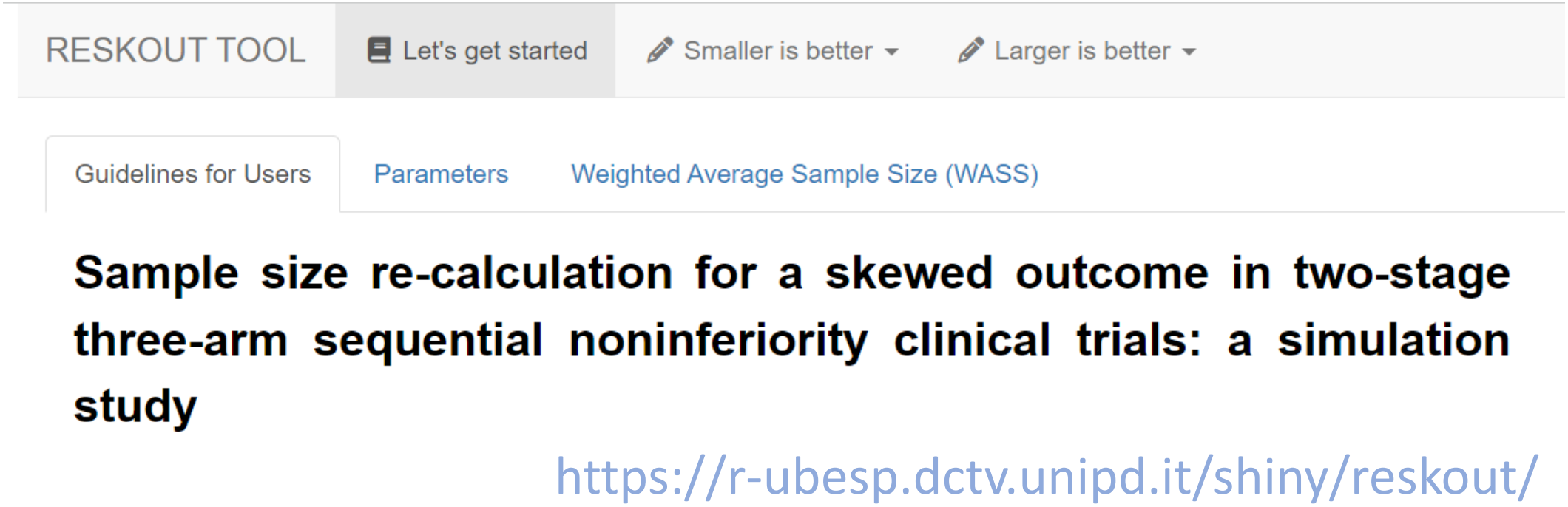
$\mu T <$ 8344 -476.5=7867.5

A real data example (3)

Groups	Total cost	
	Mean	**Standard Deviation**
Usual Care Strategy= control group	8344	8894
Early Discharge=active comparator group	7391	7903
New strategy= experimental group	7666	7488

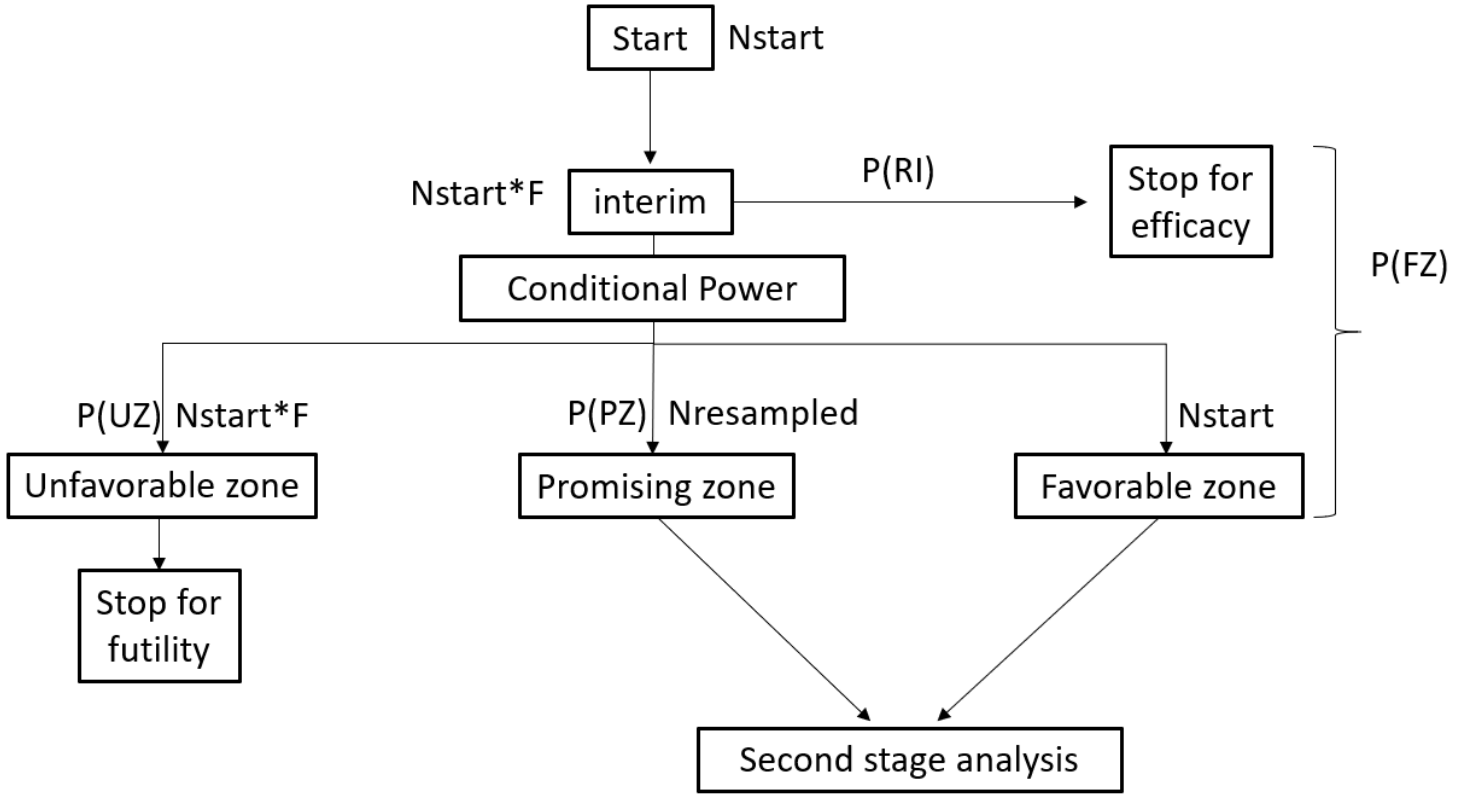
A brief presentation of the r-shiny web application

Resampling algorithm for SKewed OUTcome



https://r-ubesp.dctv.unipd.it/shiny/reskout/

Diagram of the resampling algorithm & the weighted average sample size (WASS)



$$(Nstart * F) * P(RI) + Nstart * \big(P(FZ) - P(RI)\big) + (Nstart * F) * P(UZ) + Nresampled * P(PZ)$$

Results

- Comparison between study designs under diverse hypothesis settings
- Study design by original and proposed algorithm by the real-data example

Comparison between study designs under diverse hypothesis (1)

Estimates under SD=SDP		Estimates under SDP>SDA by the proposed method									
0%		**10%**		**20%**		**30%**		**40%**		**50%**	
N	**Nr**	**N**	**Nr**	**N**	**Nr**	**N**	**Nr**	**N**	**Nr**	**N**	**Nr**
172	**189**	150	166	128	142	110	122	97	109	86	97
CP	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**
0.84	**0.2**	0.84	0.17	0.85	0.16	0.85	0.16	0.85	0.14	0.85	0.13

Comparison between study designs under diverse hypothesis (2)

Sample size assuming SD=SDP	CP and P(RI) assuming SD=SDP when SDA>SDP									
	10%		20%		30%		40%		50%	
	CP	P(RI)	CP	P(RI)	CP	P(RI)	CP	P(RI)	CP	P(RI)
172	0.77	0.12	0.72	0.09	0.66	0.06	0.62	0.05	0.56	0.03

Comparison between study designs under diverse hypothesis (3)

Estimates assuming SD=pooled SD when SDA>SDP									
10%		20%		30%		40%		50%	
N	Nr	N	Nr	N	Nr	N	Nr	N	Nr
190	215	209	238	228	264	248	287	269	313
CP	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**
0.8	0.15	0.78	0.14	0.77	0.13	0.74	0.13	0.72	0.11

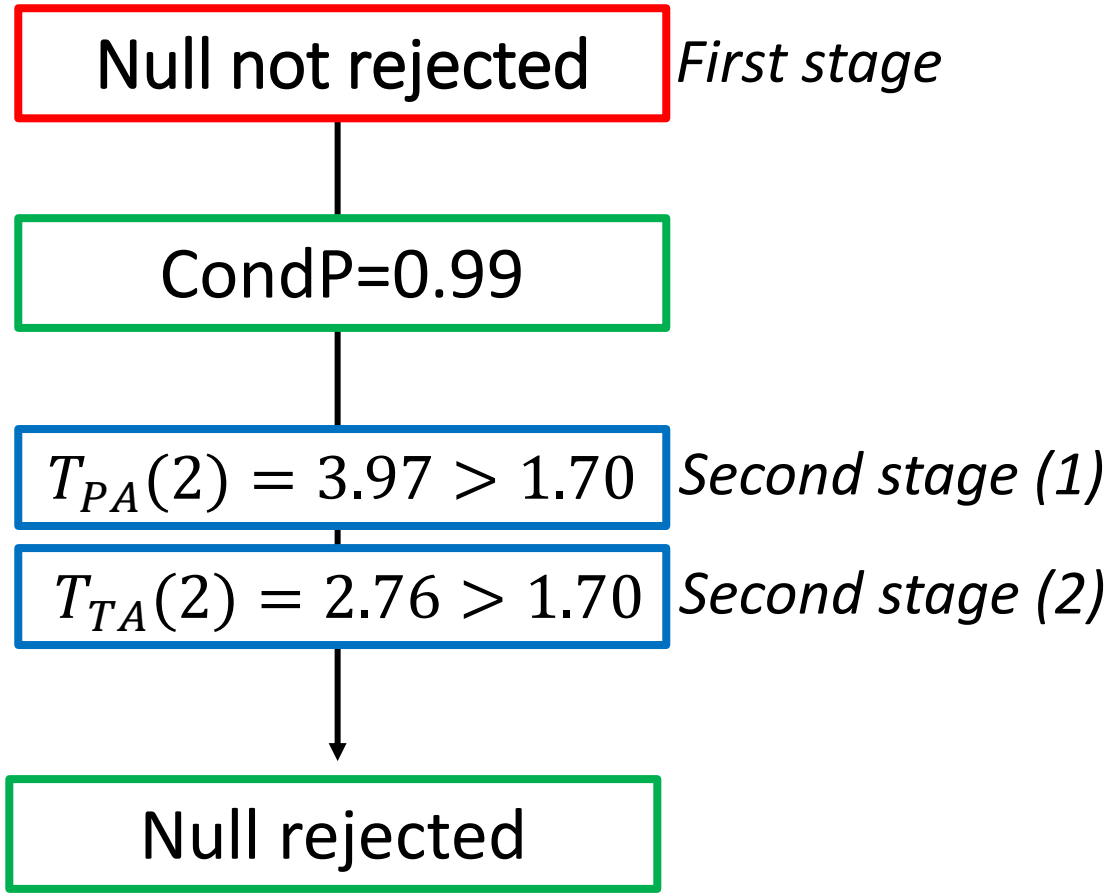
Comparison between study designs under diverse hypothesis (4)

Estimates assuming SDA>SDP when SDA>SDP									
10%		20%		30%		40%		50%	
N	Nr	N	Nr	N	Nr	N	Nr	N	Nr
214	232	252	278	294	321	339	370	388	421
CP	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**	**CP**	**P(RI)**
0.85	0.19	0.84	0.19	0.85	0.21	0.85	0.22	0.85	0.21

The real-data example (1)

Labels	original	proposed	Difference (proposed - original)
Initial sample size	3276	2730	-546
Coverage Probability	0.876	0.811	-0.065
Average resampled sample size	3628	3191	-437
Probability to reject H0 at interim	0.112	0.062	-0.05
Probability to fall in unfavorable zone	0.164	0.206	0.042
Probability to fall in promising zone	0.209	0.231	0.022
Probability to fall in favorable zone	0.627	0.563	-0.064
WASS	2897	2471	**-427**

The real-data example (2)



Conclusions

- Take home messages
- Tutorial for the r-shiny web application

Take home message (1)

To our knowledge, no one has developed a parametrical multiple testing procedure to estimate the sample size for a three-arm noninferiority study with interim analysis for skewed/gamma distributed outcomes

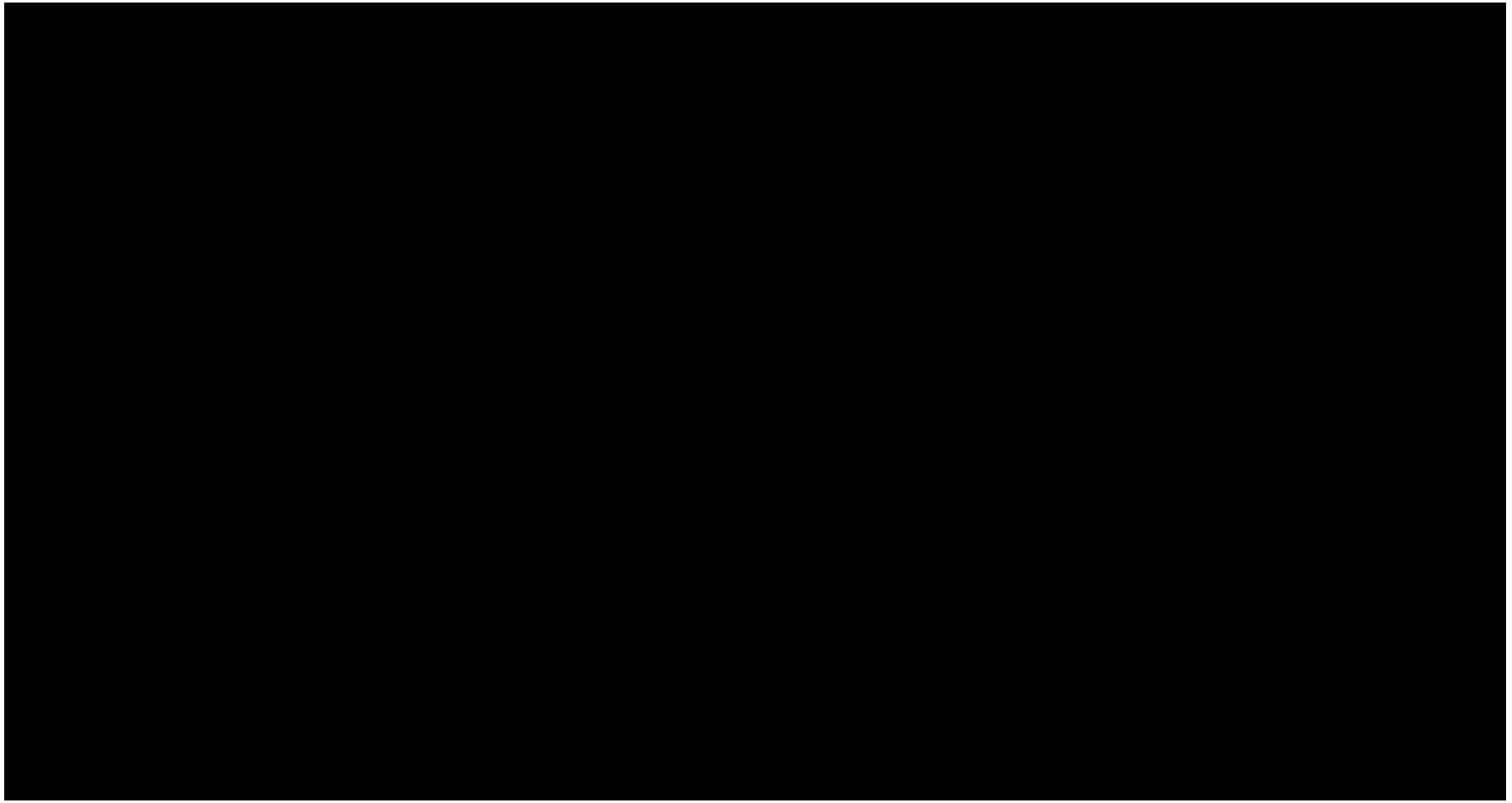
We have used a **simulation approach** to empirically understand the **properties** of the statistical test under **deviation from normality**

Take home message (2)

The first result was that **the original method**, **even under the assumption of non-normality**, **does not lose statistical power**, **assuming** the **variance** of the three groups is **equal** or very similar

We found that **if the discrepancy between the variances of the groups is considered** , using the proposed algorithm, we can estimate an initial sample size to achieve the desired power, **avoiding** the risk of **overestimation (saving patients)** or **underestimation (saving power),** even in case of deviation from normality

Tutorial for the r-shiny web application



Thank you very much for the attention !

Do you have any question?

The base-point of this study: the Lanyu Lei sample size algorithm

A group sequential design based on the promising zone of conditional power (2)

Table 4. The threshold of promising zones for the three-arm non-inferiority test in the two-stage SSR designs

n_{max} (folds of original n)	Fraction of sample size at interim look ($n^{(1)}/n^{(2)}$)	Lower bound of promising zone
2	0.5	36%
2	0.75	33%
4	0.5	31%
4	0.75	30%
∞	0.5	31%
∞	0.75	30%

Note: The lower bounds were calculated for nominal initial power of 80%.

Comparison between study designs under diverse hypothesis (5)