Mathematics & Statistics | Lancaster University

Factorial *vs* MAMS designs in clinical trials

25 September 2018

Despina Vasileiou

Outline

- Motivation and example
- Key design features
- Simulation results for different scenarios
- Case study design comparison
- Conclusions

Design comparison

Motivation

Mathematics & Statistics | Lancaster University

- The FDA call for clinical trial methods that achieved reliable results more quickly necessitates both:
  - Simultaneous study of two or more treatments within one trial.
  - Accurate estimation of the main treatment effects.
- Designs commonly used in the literature.
  - Factorial designs, which include sole treatments and their combinations.
  - Multi-Arm designs (MA).
  - Multi-Arm Multi-Stage designs (MAMS).

Common design basis

Mathematics & Statistics | Lancaster University

- Comparing two arms, A and B, and their combination, AB, against control
- Normally distributed response, $Y_j \sim \mathcal{N}(\mu_j, \sigma^2)$ with $j = A, B, AB, 0$ and μ_j the mean effect of the response to treatment or control
- Global null hypothesis testing with overall type-I error control: $H_0 = \{H_{0A} : \mu_A \leq \mu_0,\ H_{0B} : \mu_B \leq \mu_0,\ H_{0AB} : \mu_{AB} \leq \mu_0\}$
- Allocation ratios r, q for the single treatment and combination groups respectively, relative to the control group, *i.e* $n_A = n_B = rn_0$ and $n_{AB} = qn_0$, with comparisons based on balanced designs when $r = q$.

Multi-arm design features

Mathematics & Statistics | Lancaster University

- A four-arm design using Dunnett's test [2] uses the full model for treatment effect estimation.

$$Y_i = \beta_0 + \beta_1 \mathrm{I}_{Ai} + \beta_2 \mathrm{I}_{Bi} + \beta_3 \mathrm{I}_{Ai}\mathrm{I}_{Bi} + \varepsilon_i \text{ with } i = 1, 2, \ldots, n_0 + n_A + n_B + n_{AB} \quad (1)$$

Mean treatment response

Treatments		B	
		Presence	Absence
A	Presence	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_0 + \beta_1$
	Absence	$\beta_0 + \beta_2$	β_0

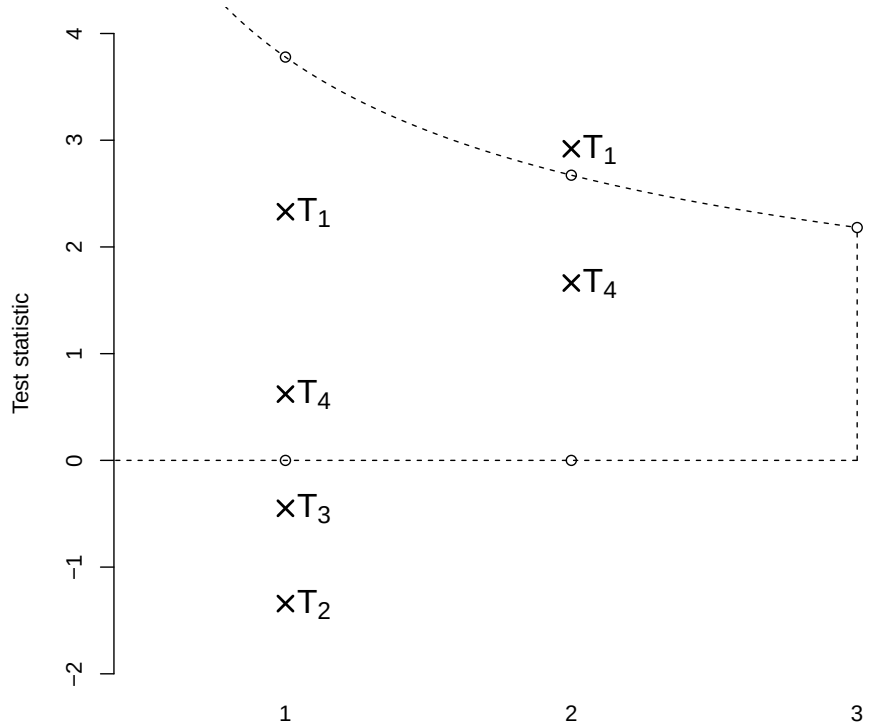
- The statistics for the hypothesis testing are based on

$$Z_{\mathbf{c}_j^\top \beta} = \frac{\mathbf{c}_j^\top \hat{\beta}}{\sigma \sqrt{\mathbf{c}_j^\top (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \mathbf{c}_j}} \quad (2)$$

Multi-arm multi-stage design features

Mathematics & Statistics | Lancaster University

- Extends Dunnett test to allow for interim analyses [3]
- Use of the O'Brien-Fleming boundary shape
- Allows for early stopping based on benefit or lack thereof
- Selects treatments that look promising



Factorial design features

Mathematics & Statistics | Lancaster University

Treatment effect and allocation impact

- Factorial designs assume no interaction in treatment effect estimation, *i.e.* $\beta_3 = 0$. In a 2×2 design:

$$Y_i = \beta_0 + \beta_1 \mathrm{I}_{Ai} + \beta_2 \mathrm{I}_{Bi} + \varepsilon_i$$

- The test statistics used are

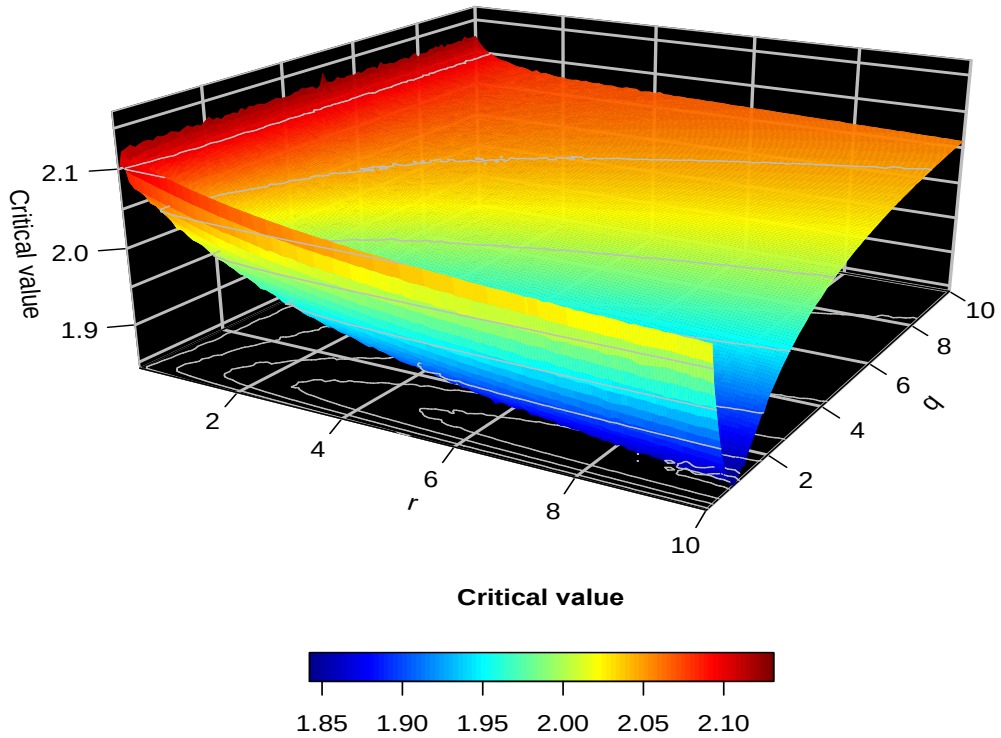
$$\begin{aligned}
Z_A &= \sqrt{n_0}\frac{r(r+q)(\bar{Y}_A - \bar{Y}_0) + qr(1+r)(\bar{Y}_{AB} - \bar{Y}_B)}{\sigma\sqrt{(1+r)(r+q)(r^2+2rq+r^2q)}} \\
Z_B &= \sqrt{n_0}\frac{r(r+q)(\bar{Y}_B - \bar{Y}_0) + qr(1+r)(\bar{Y}_{AB} - \bar{Y}_A)}{\sigma\sqrt{(1+r)(r+q)(r^2+2rq+r^2q)}} \\
Z_{AB} &= \sqrt{n_0}\frac{\bar{Y}_{AB} - \bar{Y}_0}{\sigma\sqrt{\frac{1+q}{q}}} \qquad (3)
\end{aligned}$$

Factorial design features

Mathematics & Statistics | Lancaster University

Allocation impact on critical values ($\alpha = 0.05$)

- Balanced design the critical value is found to be $k = 2.028$
- When $r = q$ optimal which corresponds to critical value 2.017 is for $r = q = 1.7$
- For $r \in [0.5, 2.5]$ the optimum critical value of 1.954 occurs for $q = 0.8$,$r = 2.5$



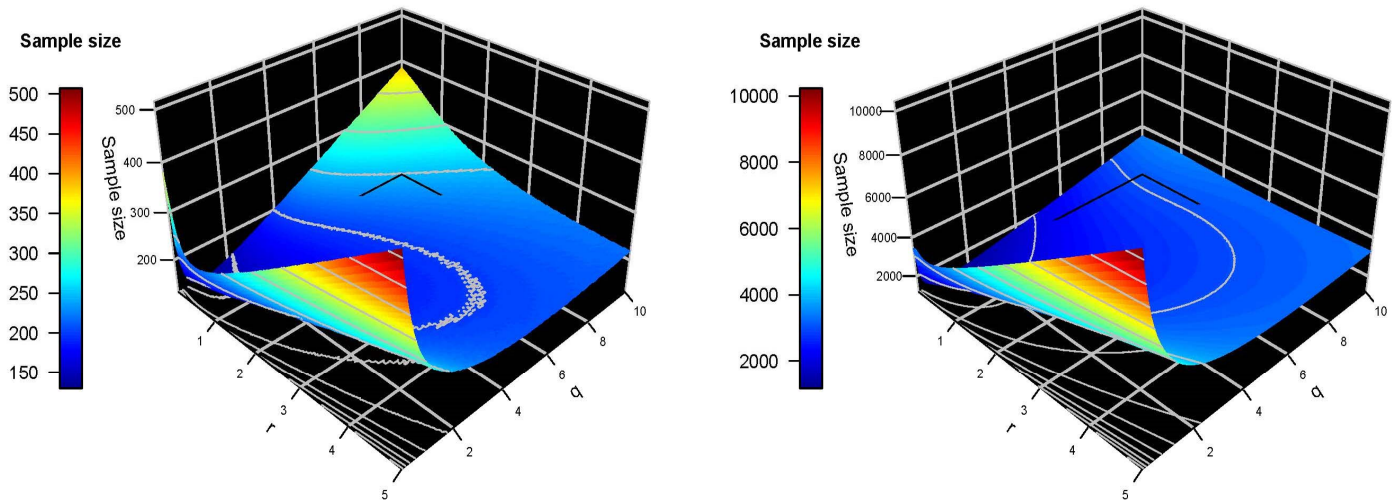
r \ q	0.1	0.5	1
0.1	2.09	2.10	2.11
0.5	2.07	2.05	2.07
1	2.06	2.02	2.03
2	2.04	1.97	1.98

Alternative scenarios

Allocation impact on sample size

Mathematics & Statistics | Lancaster University

- Alternative hypothesis scenarios consistent with factorial design assumptions with $\Delta = 0.5$ and $\delta_0 = 0.1$
  1. $\mathrm{H}_1 : \mu_A - \mu_0 = \Delta, \mu_B - \mu_0 = \delta_0, \mu_{AB} - \mu_0 = \Delta + \delta_0, \mu_0 = 0$
  2. $\mathrm{H}_1 : \mu_A - \mu_0 = \mu_B - \mu_0 = \delta_0, \mu_{AB} - \mu_0 = 2\delta_0$



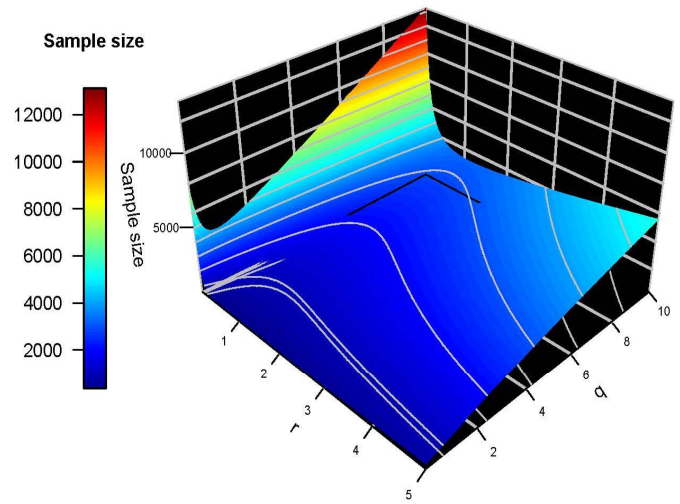
- H_1 on the left: Balanced design sample size 160 with minimum 129 when $r = 0.01$ and $q = 0.9$
- H_1 on the right: Balanced design sample size 2008 with minimum 1150 when $r = 0.01$ and $q = 1$

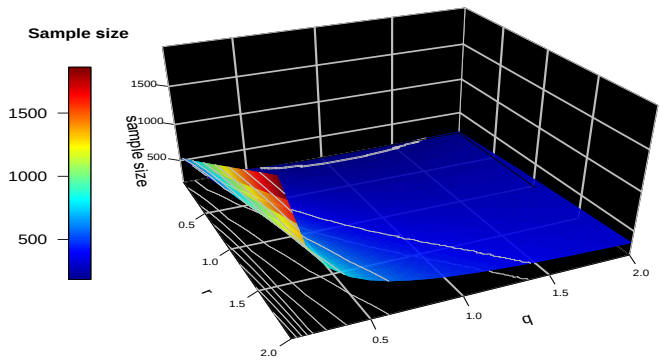
Alternative scenarios

Allocation impact on sample size

Mathematics & Statistics | Lancaster University

- Alternative hypothesis scenarios inconsistent with factorial design assumptions with $\Delta = 0.5$ and $\delta_0 = 0.1$
  1. $\mathrm{H}_1 : \mu_A - \mu_0 = \Delta$ & $\mu_B - \mu_0 = \mu_{AB} - \mu_0 = \delta_0$
  2. $\mathrm{H}_1 : \mu_{AB} - \mu_0 = \Delta, \mu_A - \mu_0 = \mu_B - \mu_0 = \delta_0$





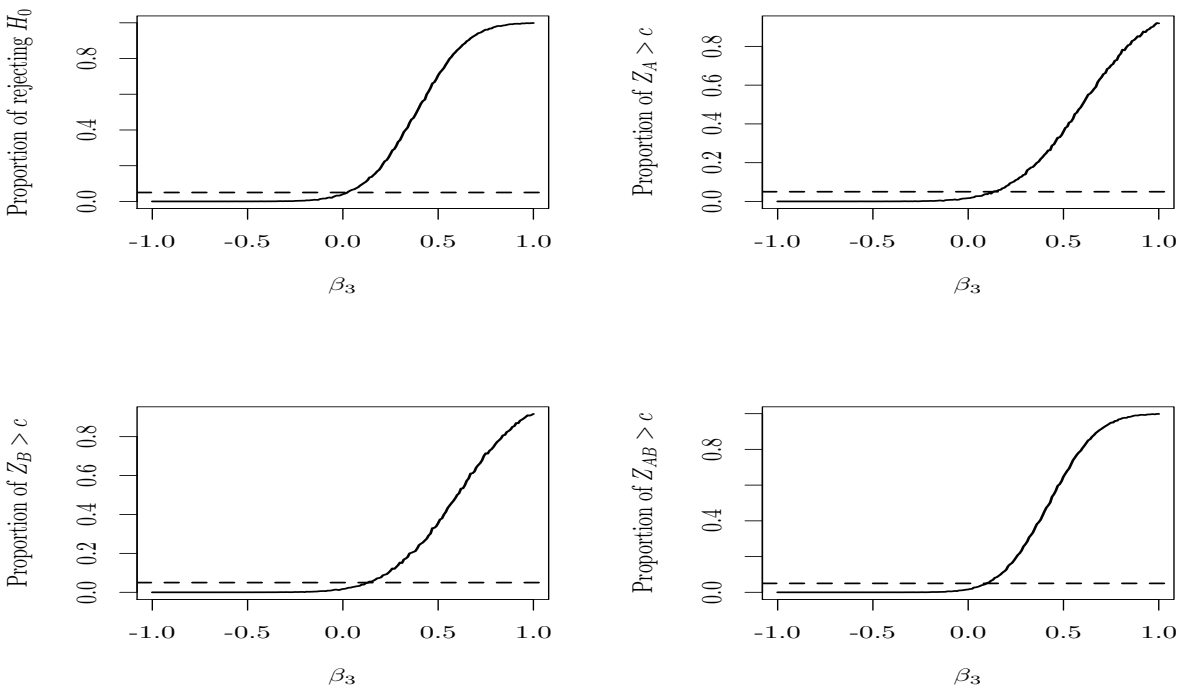
- H_1 on the left: Balanced design sample size 704 with minimum 326 when $r = 0.81$ and $q = 0.1$
- H_1 on the right : Balanced design sample size 324 with minimum 199 when $r = 0.1$ and $q = 1$

Simulation Results

Mathematics & Statistics | Lancaster University

Effect of interaction on factorial designs

- Explore additivity of treatment effects in balanced designs
- β_3, ranges from -1 to 1 (antagonism to synergy)



For $H_1 : \{H_{1A} : \mu_A > 0, \text{or } H_{1B} : \mu_B > 0, \text{or } H_{1AB} : \mu_{AB} > 0\}$, and for the remaining three plots $H_{1j} : \mu_j > 0$ for each $j = A, B, AB$.

Comparison Results

Mathematics & Statistics | Lancaster University

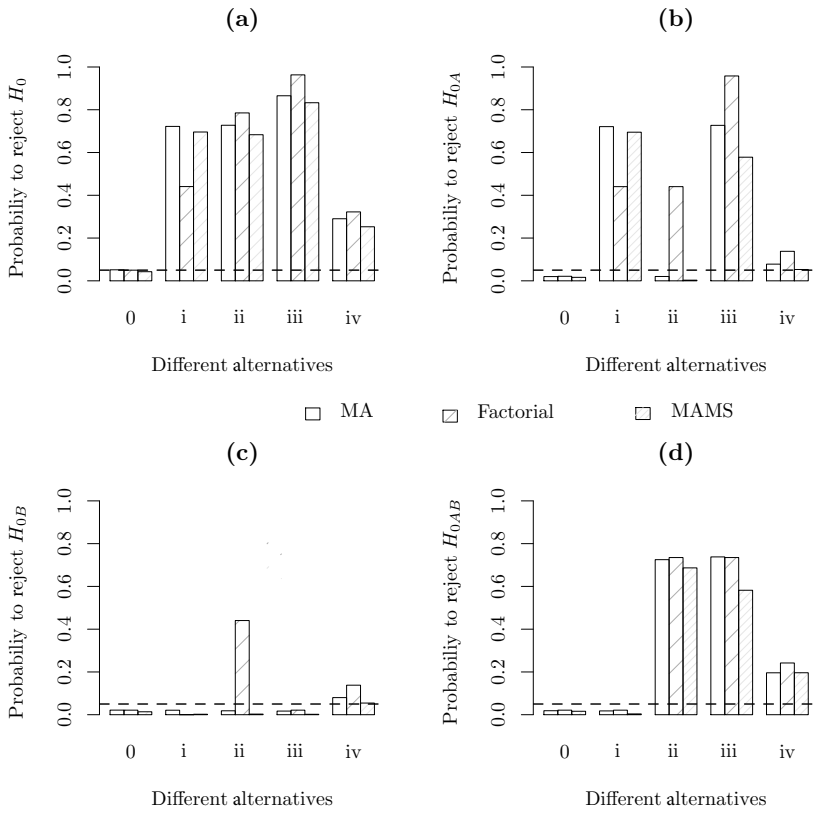
Direct power comparison between all designs

- Based on study evaluating use of physiotherapy on osteoarthritis [1]
- Either manual physiotherapy, exercise physiotherapy, both or standard of care
- $n = 45$ per group
- Difference in points of WOMAC score
- Interesting effect $\Delta = 28$, uninteresting $\delta_0 = 7$ and $\sigma = 50$.
- Performance of Factorial, MA and MAMS designs

Comparison Results

Mathematics & Statistics | Lancaster University

Probability of rejecting the null hypothesis



0: $\mu_0 = \mu_A = \mu_B = \mu_{AB} = 0$

i: $\mu_A - \mu_0 = \Delta,$
$\mu_B - \mu_0 = \mu_{AB} - \mu_0 = \delta_0$

ii: $\mu_A - \mu_0 = \mu_B - \mu_0 = \delta_0,$
$\mu_{AB} - \mu_0 = \Delta$

iii: $\mu_A - \mu_0 = \Delta, \mu_B - \mu_0 = \delta_0,$
$\mu_{AB} - \mu_0 = \Delta + \delta_0$

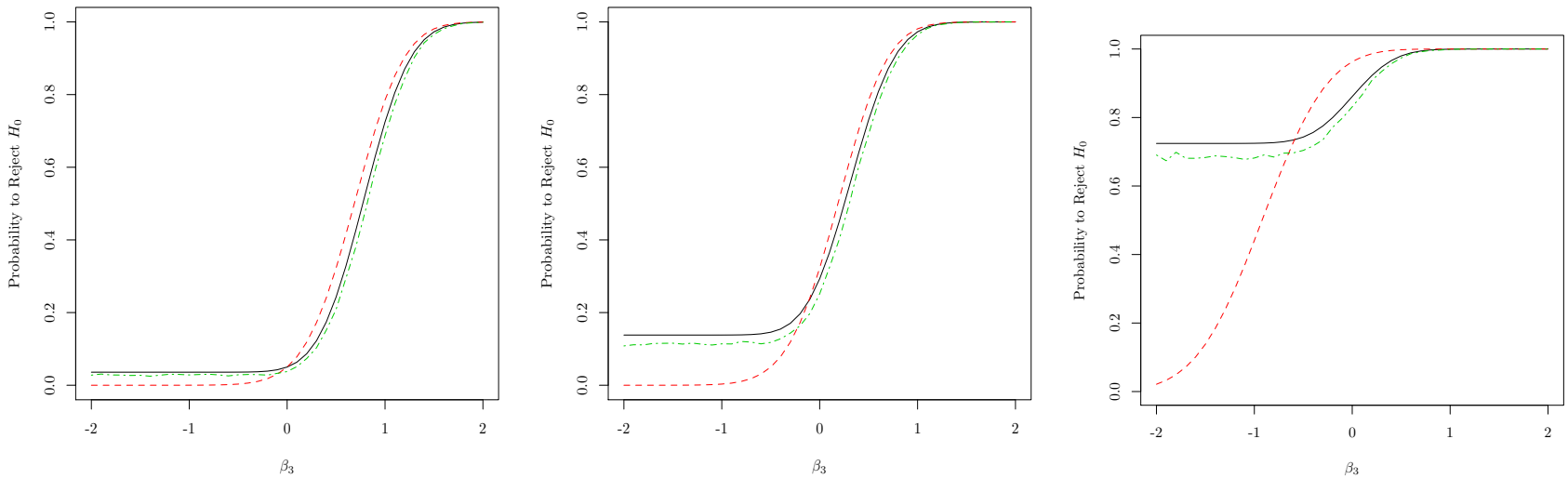
iv: $\mu_A - \mu_0 = \delta_0, \mu_B - \mu_0 = \delta_0,$
$\mu_{AB} - \mu_0 = 2\delta_0$

Interaction effect

Mathematics & Statistics | Lancaster University

Probability of rejecting the null hypothesis, while the interaction ranges from -2 to 2 ($*\Delta$) when:

1. $\mu_A - \mu_0 = \mu_B - \mu_0 = 0$
2. $\mu_A - \mu_0 = \mu_B - \mu_0 = 7$
3. $\mu_A - \mu_0 = 0$ & $\mu_B - \mu_0 = 28$.



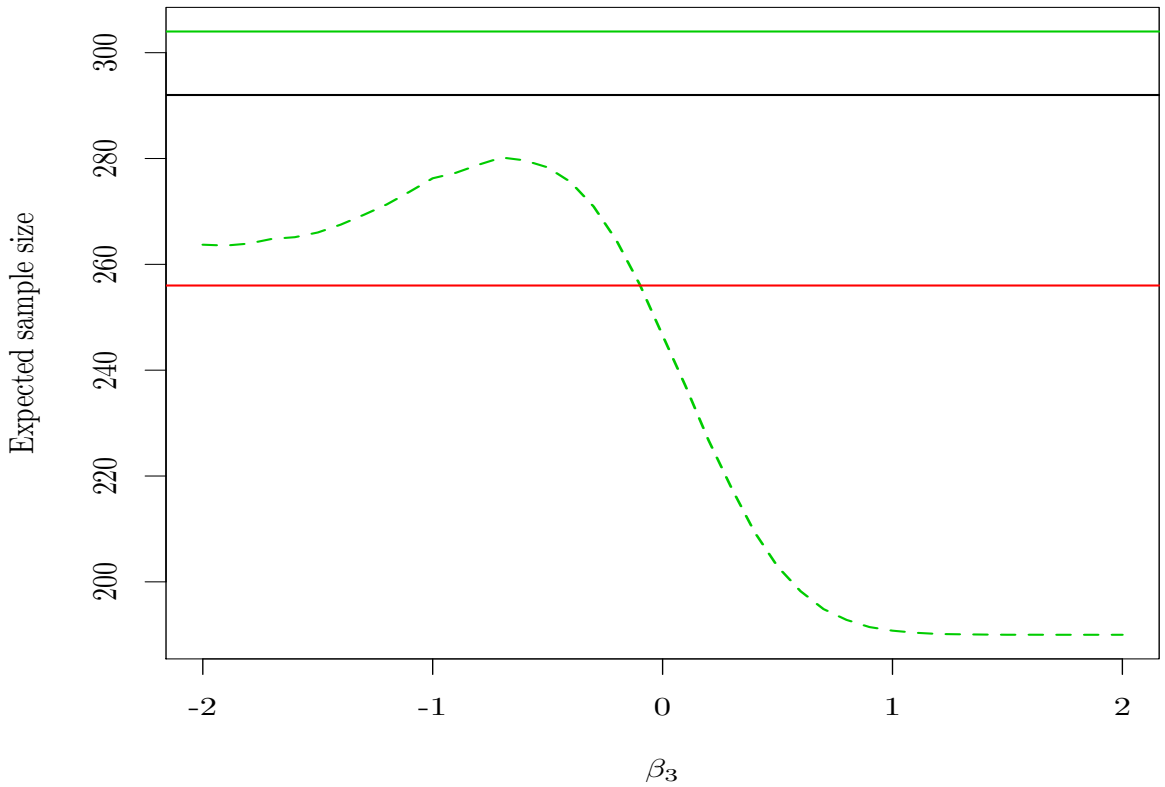
Red for factorial design, black for MA design and green for MAMS design

Design differences

Total sample size

Mathematics & Statistics | Lancaster University

Comparison amongst sample sizes of a balanced factorial design, a multi-arm design and the expected sample size of a multi-arm two-stage design with 0 futility boundary using case study parameters and $\alpha = 0.05$, $1 - \beta = 0.9$.



Conclusions

- No difference in the expected sample size of a MAMS trial and a factorial one when there is no interaction between the treatments.
- Observed a notable inflation of the type I error in the simulation study when the sole treatments interact in a synergistic manner ($\beta_3 > 0$).
- Also found losses of power when the treatments have in combination an antagonistic effect.
- Factorial designs should only be considered instead of a multi-arm design when there is evidence that the assumption of additivity is met.
- MAMS designs are a robust alternative to the presence of interactions and are expected to require a much smaller sample size at the expense of a small deficiency in power.

References

Mathematics & Statistics | Lancaster University

J. H. Abbott, M. C. Robertson, J. E. McKenzie, G. D. Baxter, J.-C. Theis, A. J. Campbell, et al.

Exercise therapy, manual therapy, or both, for osteoarthritis of the hip or knee: a factorial randomised controlled trial protocol.
Trials, 10(11):6215–10, 2009.

C. W. Dunnett.

Selection of the best treatment in comparison to a control with an application to a medical trial.
Design of experiments: Ranking and selection, pages 47–66, 1984.

D. Magirr, T. Jaki, and J. Whitehead.

A generalized dunnett test for multi-arm multi-stage clinical studies with treatment selection.
Biometrika, 99(2):494–501, 2012.

Published Paper

Mathematics & Statistics | Lancaster University

Jaki, T. and Vasileiou, D.
Factorial versus multiarm multistage designs for clinical trials with multiple treatments.
Statistics in medicine, 36(4), pp.563-580,2017.

Mathematics & Statistics | Lancaster University

25 September 2018

Despina Vasileiou

Medical and Pharmaceutical
Statistics Research Unit