Achieving Robust Evidence while Minimising Expected Control Group Size - A Bayesian Adaptive Trials Perspective

Sof´ıa S. Villar

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Forum PBC Meeting

Edinburgh, April 7

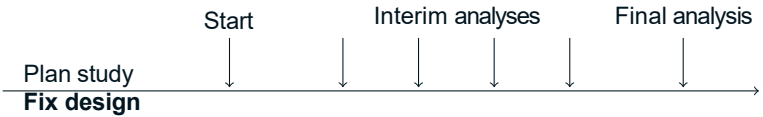
Disclosure: Current Scientific Advisor to Phase V and Former Visiting Researcher at Google.

Outline

Fixed Sample Design

- **total** sample size known in advance
- no adjustment possible
- Requires strong assumptions in the planning stage

Adaptive Design



At each interim example of possible adaptations include:

- decide whether or not to stop
- change the target total sample size
- change the allocation ratio per treatment groups
- change the endpoint
- ... more

The additional flexibility results in higher variability of the resulting design and this can result in e.g.:

- lower **expected** sample size (advantage of the flexibility)
- larger **maximum** sample size (disadvantage to preserve integrity of analysis)

Outline

The ARREST Trial - The Condition and Motivation

- More than **350,000 people** in the US **die** from out-of-hospital cardiac arrest (**OHCA**) each year.
- Patients with refractory ventricular fibrillation/pulseless ventricular tachycardia (VF/VT) OHCA have one of the worst prognosis (**85-90% mortality**), however the **underlying cause** of cardiac arrest is likely to be **reversible**.
- Adult patients (18-75 years old) with VF and VT who are transferred by emergency medical services (EMS) will receive one of the **2 standards of care practiced**.
  - Extracorporeal membrane oxygenation facilitated resuscitation **(ECMO) (Treatment** 1**)**
  - Standard advanced cardiac life support resuscitation **(ACLS) (Treatment** 0**)**.
- The primary objective: **Survival to hospital discharge**.
- The primary endpoint is a **binary endpoint** observable in 30 days.

The ARREST Trial - Hypotheses

- Phase II, single center, partially blinded, intention to treat, safety and efficacy clinical trial where **primary study hypotheses** are

$$H_0 : p_1 = p_0 \text{ vs } H_A : p_1 \neq p_0.$$

  - p_1 the **probability** of a **positive response** in the target population under ECMO.
  - p_0 the **probability** of a **positive response** in the target population under ACLS.

- We are **interested in** $P(p_1 > p_0|Data) = 1 - P(p_1 < p_0|Data)$.

The ARREST Trial - Priors and Posteriors

- *Beta*(1, 1) **prior distributions** for p_1 and p_0.
- *Beta*($s_1 + 1, n_1 - s_1 + 1$) **posterior distribution** for p_1.
- *Beta*($s_0 + 1, n_0 - s_0 + 1$) **posterior distribution** for p_0.
  - n_1, n_0 denotes the number of participants assigned to treatment 1 and 0.
  - s_1, s_0 denotes the number of positive responses (successes) observed in treatment 1 and 0.
- We then calculate

$$P(p_1 > p_0 | Data)$$

- In a Bayesian adaptive design, this probability will inform interim decisions.

The ARREST Trial - Early Stopping

- Interim analyses are **conducted after every 30 patients are randomised**.
- The trial will stop for **efficacy** if:

$$P(p_1 > p_0 | Data) \geq \xi.$$

  - ξ is the level needed to reject the **null hypothesis**.
- The trial will stop for **inferiority** if:

$$P(p_1 < p_0 | Data) \geq \xi.$$

The ARREST Trial - Response Adaptive Randomisation

- The initial **burn-in is size 30** with equal allocation (15 patients in each treatment).
- If the trial does not stop, a form of Bayesian Response-adaptive randomisation is used. The next group of patients is randomized to ECMO with a target probability that is proportional to its posterior probability of superiority , i.e., $P(p_1 > p_0|Data)$.
- An additional restriction is the **randomization probability may not exceed 75%** in either direction.

The ARREST Trial - Type I Error and Power

- Aims: **Type 1 error** rate of 5% two sided, and **power** of 90%, assuming success rates of 12% vs. 37% in the 2 groups.

- This was achieved by a design with $\xi = 0.986$ and up to **5 analyses interim after 30 patients** in an **expected required sample size** is $N = 148$ (inflated to N = 174 for **15% expected drop out**). Only the **first 150 patients** are evaluated.

	Scenario	Prob reject null	$E(N)$	$E(N_1)$	$E(N_0)$	n- $E(N)$	n/2- $E(N_0)$
GS-RAR	Null	0.048	148.5	74.2	74.3	-12	-6
GS-RAR	Alternative	0.905	81.6	52.5	29.2	54	39
GS-ER	Null	0.050	148.8	74.4	74.4	-13	-6
GS-ER	Alternative	0.930	83.2	41.6	41.6	53	26
ER	Null	0.050	136	68	68	0	0
ER	Alternative	0.950	136	68	68	0	0

Table 1: Simulated operating characteristics of two adaptive designs against the non-adaptive RCT.

- $E(N)$ expected total sample size, $E(N_1)$ expected sample size for ECMO and $E(N_0)$

Outline

Re-use of Historical Data in Clinical Trials

- There are lots of data available (particularly on control treatments). How can we use these to reduce the size / risk of failure a new trial by re-using that data?
- Historical controls are common in device and early oncology trials. Current interest lies in paediatric drug development (partial adult efficacy/safety extrapolation) and historical controls for late-phase/regulatory use.
- Most approaches for handling historical data downweight the historical information. When we maximise the downweight, no historical data are used in the current study, if we maximise the borrowing/learning from it then all the historical data are used in a pooled analysis with current data (no downwweight at all of past data).
- How to determine this weight? Modified power prior, empirical power prior, or equivalence probability weight approach (by Maxine Ajimi, formally known as M. Bennett)

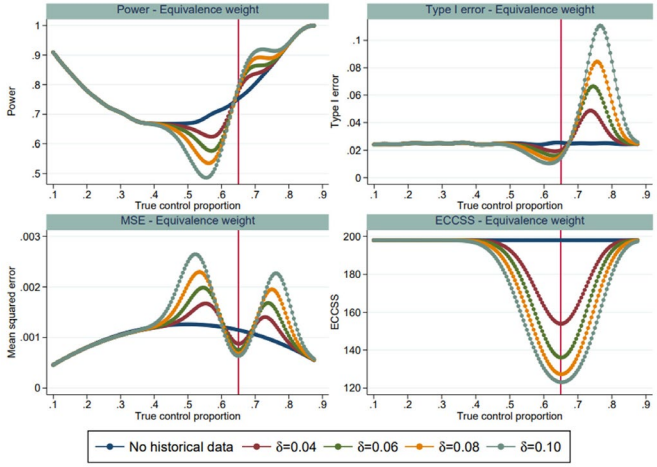
Adaptive partial replacement design

- Standard two arm RCT where 400 patients (randomised 1:1), 12% treatment effect (TE) to detect, 2.5% error rate and 76% power. Binary endpoint (success/failure)
- Historical data available on 100 (n_h) control patients with a response probability of 65%.
- Adaptive design (changing allocation ratio)
  - Stage 1: Randomise n_t^1 to treatment and n_c^1 to control (e.g. $n_t^1 = n_c^1 = 100$); Calculate the weight for historical data using the first stage controls and the historical data. The resulting weight determines $ESSh1$.
  - Stage 2: Randomise $(n_t - n_t^1)$ to treatment and $\max(n_c - n_c^1 - ESSh1; nmin)$ to control (e.g., $nmin = 20$). Recalculate the weight, at the end of the at the end of the study using all current control data. Compute a posterior probability for the final decision rule $P(TE \geq 12\%|data) > \xi$

The equivalence approach is flexible and the equivalence bounds can be chosen to control the maximum inflation in type I error across all true current control proportions.

Operating Characteristics

When perfect alignment, with most/least conservative equivalence criteria the reduction in control group is of 40 (200-160)/ 75 (200-125).



Outline

Adaptive Designs in Rare Diseases

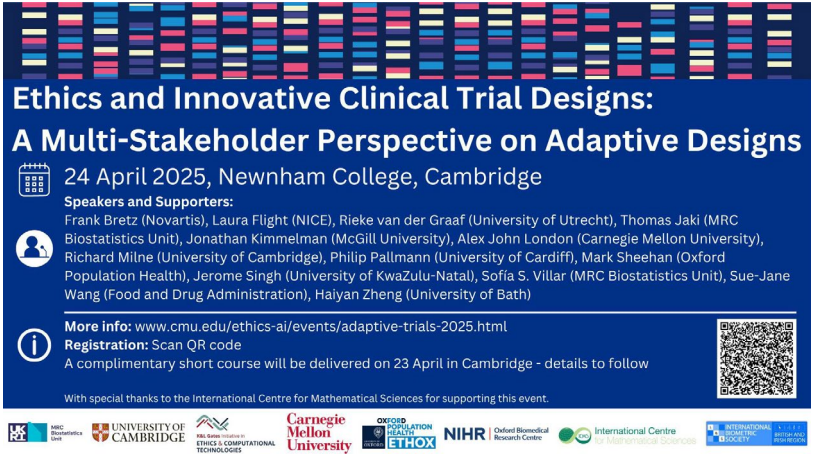
In many settings, designing trials to meet conventional evidentiary standards may not be feasible → Balance is needed between what is necessary and what is possible.

To achieve this with an adaptive design in the PBC setting it is essential to consider:

- The validation of a quickly observable endpoint (could be continuous).
- The use of a model that combines data from short and longer term endpoints to make interim decisions.
- The possibility of a platform trial with a shared placebo arm can also imply a reduction in the placebo arm without impacting power.
- Analysis methods crucially determined to give type I error guarantees.

Acknowledgments

Thank you for listening! :)



References

Adaptive Designs Burnett, T., Mozgunov, P., Pallmann, P. et al. Adding flexibility to clinical trial designs: an example-based guide to the practical use of adaptive designs. BMC Med 18, 352 (2020). https://doi.org/10.1186/s12916-020-01808-2

ARREST trial

Yannopoulos, D and others (2020) Advanced REperfusion STrategies for Refractory Cardiac Arrest (The ARREST Trial) Protocol version 1.4 (2020a)

Yannopoulos, D and others (2020b) Rationale and methods of the Advanced R2Eperfusion STrategies for refractory cardiac arrest (ARREST) trial. *American Heart Journal*. 229 29–39.

Yannopoulos, D and others (2020c) Advanced reperfusion strategies for patients with out-of-hospital cardiac arrest and refractory ventricular fibrillation (ARREST): a phase 2, single centre, open-label, randomised controlled trial. *The lancet*. 396 1807–1816.

Equivalence Weighted Use of Historical Data

Bennett, M. and White, S. and Best, N. and Mander, A. (2021) A novel equivalence probability weighted power prior for using historical control data in an adaptive clinical trial design: A comparison to standard methods *Pharmaceutical statistics 20(3)462–484 Wiley Online Library*

Ensuring intergrity robustly

Regulators do not tarde-off type I and tye II errors

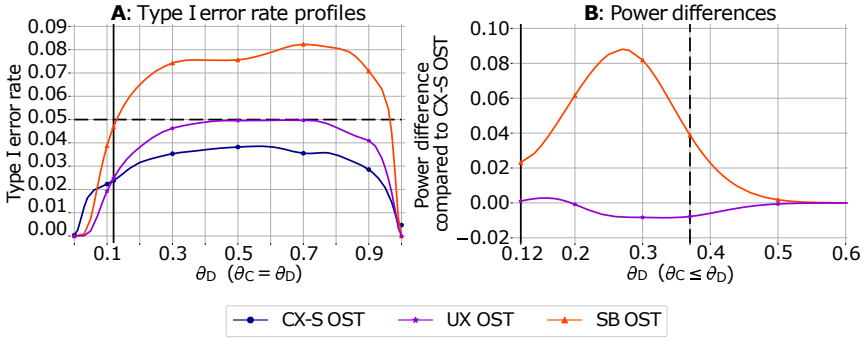


Figure 2: ARREST trial. Type I error under the SB (simulation based approach), CX-S (Conditional exact), and UX OST (Unconditional exact) for $p_0 = p_1$ (left), vertical line at $p_0 = p_1 = 0.12$. Power difference (right) for two tests compared to the CX-S OST vertical line denotes $p_1 = 0.37$.
https://arxiv.org/abs/2407.01055v2