A multi-arm multi-stage clinical trial design for binary outcomes with application to tuberculosis

Daniel Bratton

MRC Clinical Trials Unit at UCL

20th May 2014

Outline

- The need for novel trial designs
- Multi-arm multi-stage (MAMS) designs for time-to-event outcomes
- Extension to binary outcomes
- Feasible MAMS designs
- Admissible MAMS designs
- Application to tuberculosis (TB)

The need for novel trial designs

- Despite the increasing pace of drug discovery, there is a slowdown in the rate at which new therapies reach patients
- A major drawback is the inefficiency and cost of the conventional approach of testing new therapies in separate controlled trials
- If there are multiple treatments to assess a more efficient approach might be to...
  - ➢ evaluate multiple new regimens in a single trial
  - ➢ prematurely stop recruitment to poorly performing arms
  - ➢ incorporate phases 2 and 3 into a single, seamless trial

Multi-arm multi-stage (MAMS) designs

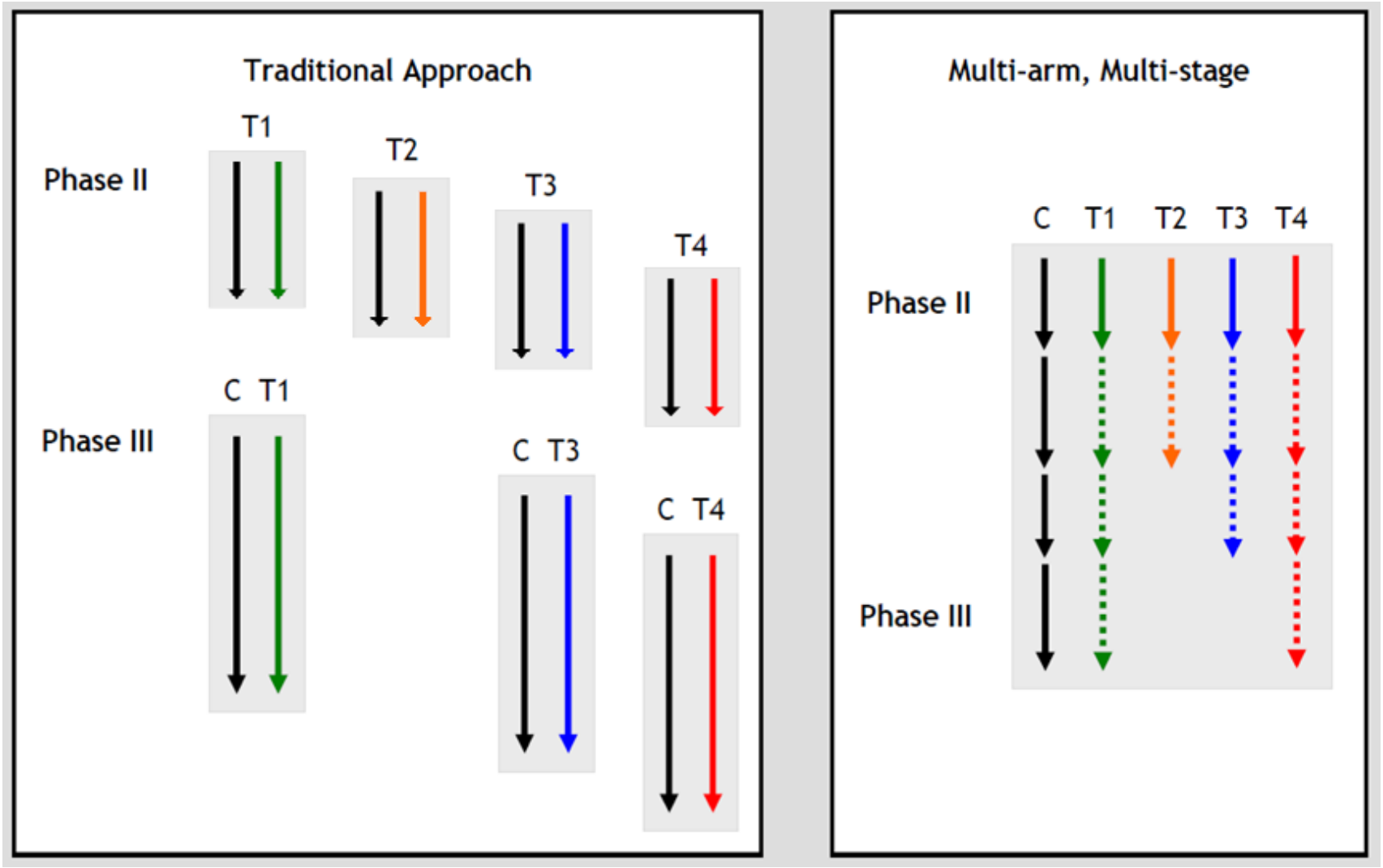
- A MAMS design which achieves these aims was developed by Royston and colleagues (2003 & 2011) for time-to-event outcomes and has been used to design several oncology trials

- Each stage can be considered as a conventional multi-arm trial with its own one-sided significance level and power for each pairwise comparison

- Interim assessments can be made on an intermediate outcome (I) which is on the causal pathway to the definitive, final outcome (D) of the trial

- E.g. I = phase 2 outcome, D = phase 3 outcome

Traditional vs. MAMS trial designs



Example – the STAMPEDE trial

- 6-arm 4-stage STAMPEDE trial in prostate cancer
- I = progression free survival (PFS), D = overall survival (OS)

Stage	Target HR	Outcome measure	One-sided sig. level	Power	Control events	Critical HR
1	0.75	PFS	0.500	0.95	113	1.00
2	0.75	PFS	0.250	0.95	216	0.92
3	0.75	PFS	0.100	0.95	334	0.89
4	0.75	OS	0.025	0.90	403	0.84
Overall			0.025	0.85		

- Overall pairwise type I error rate is bounded above by the final stage significance level since an arm may have a true effect on I but not on D
- Overall power is calculated by P(pass all stages | H_1) accounting for the between-stage correlation

MAMS design for binary outcomes

- Original MAMS design only allowed the use of time to event outcomes
- To make it more widely applicable to other areas we extended the design to binary outcomes observed a fixed time after randomisation, e.g. treatment success/failure at 1 month
- This design uses the same principles as the original design but calculation of sample size and between-stage correlation differs
- Standard formulae can be used to calculated required sample size for each stage
- If I≠D an estimate of the positive predictive value P(D=1|I=1) is required (e.g. from previous data) to estimate correlation structure

Example – 2-arm 2-stage designs (I=D)

- Outcome = treatment success at 1 month
- Control event rate = 70%
- Target difference under H_1 = 15%
- 1:1 allocation ratio

Stage (j)	α_j	ω_j	N_j
1	0.30	0.95	142
2	0.025	0.90	316
Overall	α = 0.0228	ω = 0.877	

- NB: sample size for fixed sample (1-stage) design with α = 0.0228 and ω = 0.877 is 310

How do we design a MAMS trial?

Stage	1	2	...	J	Overall
Sig. level	α_1	α_2	...	α_J	α
Power	ω_1	ω_2	...	ω_J	ω

- Need to find stagewise parameters $\{\alpha_j, \omega_j\}$ such that trial has a pre-specified overall type I error rate α and power ω (feasible design)
- Use a search procedure over $2J$ parameters $\alpha_1, ..., \alpha_J, \omega_1, ..., \omega_J$ with constraints to decrease search time
- Many feasible designs are likely to exist – which do we use in practice?

Admissible designs

- Those which minimise the expected sample size E(N) under a certain scenario are most appealing
- Common choices are minimax design (lowest max(N)) or null-optimal design (lowest $E(N|H_0)$)
- However, these can perform poorly under effects for which they are not optimised
- Instead, other designs which minimise a weighted sum of these two measures

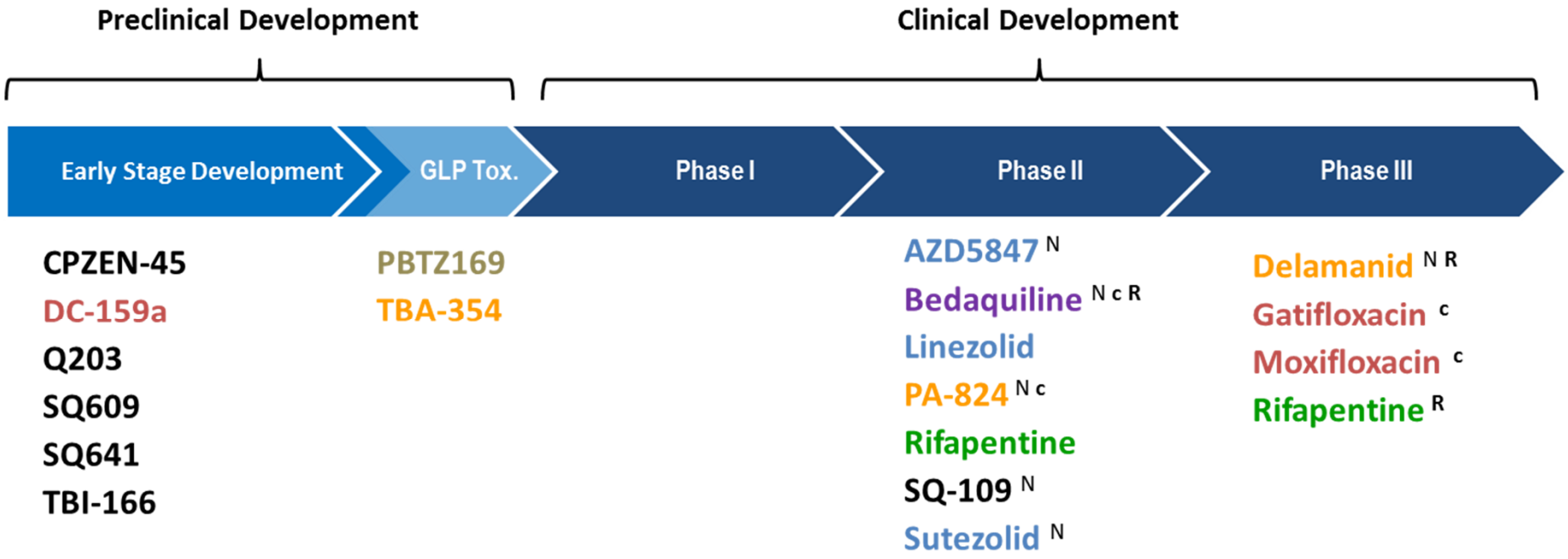
$$L = wE(N|H_0) + (1-w)\max(N) \qquad w \in [0,1]$$

known as *admissible* designs (Jung et al. 2004) are likely to be more appealing in practice

Application to TB

Global TB drug pipeline

MRC | Clinical Trials Unit



Chemical classes: fluoroquinolone, rifamycin, oxazolidinone, nitroimidazole, diarylquinoline, benzothiazinone

[1] Details for projects listed can be found at http://www.newtbdrugs.org/pipeline.php and ongoing projects without a lead compound series identified can be viewed at http://www.newtbdrugs.org/pipeline-discovery.php.

[C] Drug candidate currently in combination regimen in clinical testing

[R] Submitted for approval or approved by stringent regulatory authority (i.e., FDA, EMA, WHO Prequalification)

[N] New chemical entity

MRC | Clinical Trials Unit

WORKING GROUP ON NEW TB DRUGS

www.newtbdrugs.org

Updated: June 2013

Conventional TB trial designs

Design Parameter	Study	
	Phase II	Phase III
Primary outcome	Culture status	Failure/Relapse
Length of follow-up	8 weeks*	18 months*
Significance level (1-sided)	0.025	0.025
Power	80%	85%
Control event rate	75%	10%
Null treatment effect	0%	6% (NI margin)
Target treatment effect	13%	0%
Allocation ratio (E:C)	1:1	1:1
Attrition rate	15%	20%

*Plus an additional 6 weeks, typically, to determine culture status

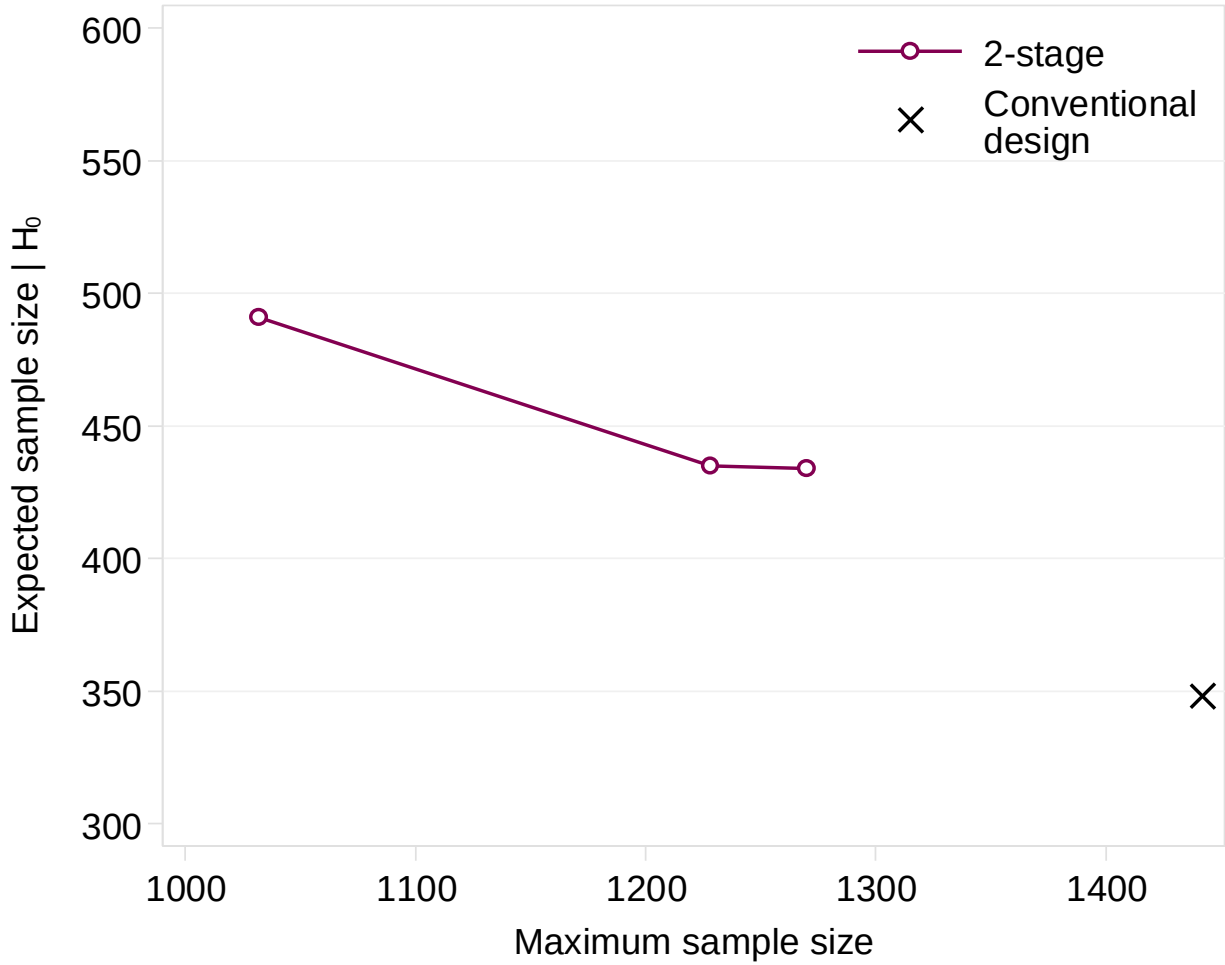
- **Required sample sizes are 320 for phase II and 1122 for phase III**
- **Total power over both phases = 68%**
- **Expected sample size under H_0 = 348**

Two-arm two-stage phase II/III TB designs

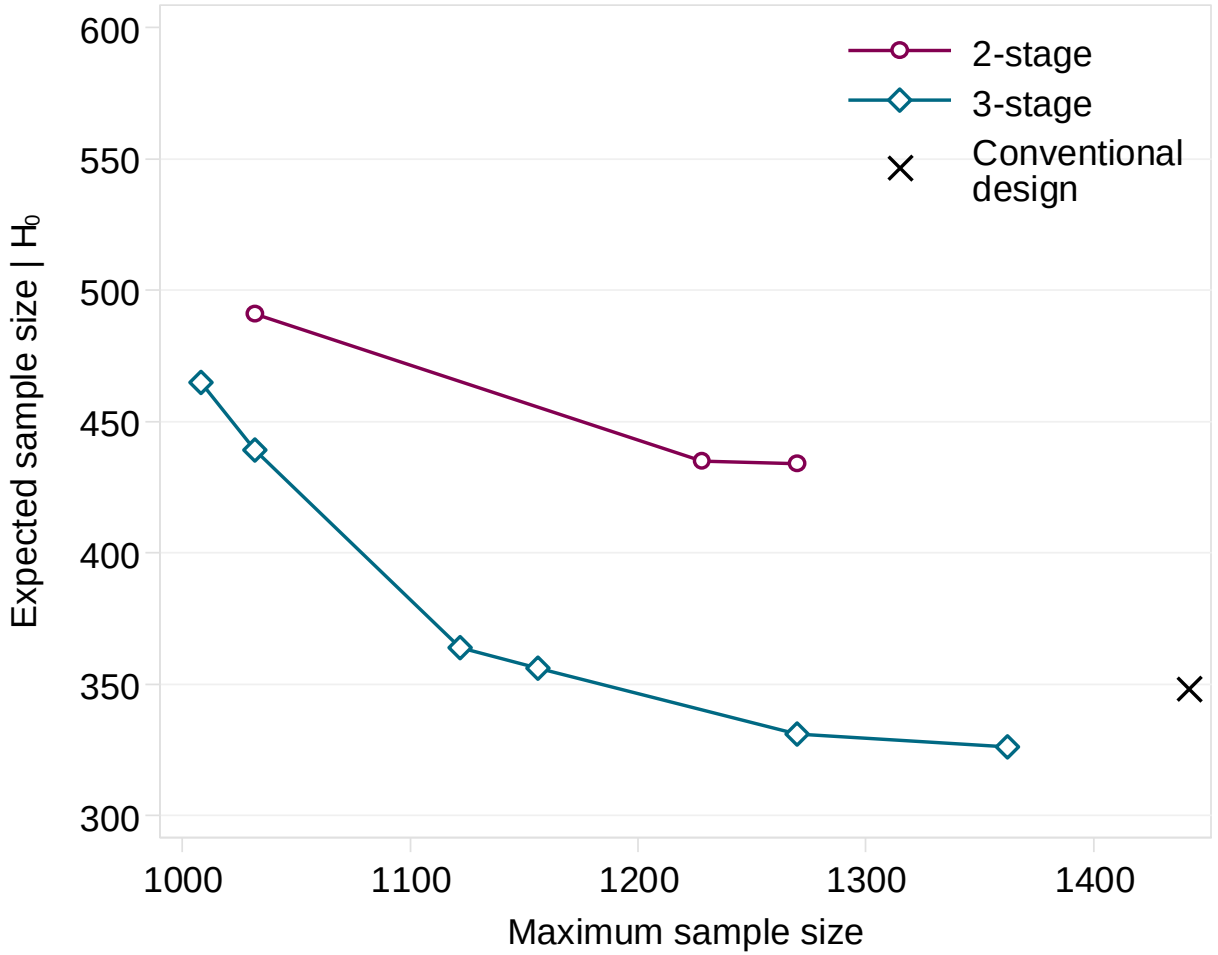
Design	Stage	Outcome	Significance level (1S)	Power	Sample size	$E(N \mid H_0)$
Null-optimal	1	Culture Status	0.05	0.89	390	434
	2	Relapse	0.025	0.89	1270	
		Overall	0.025	0.80		
Minimax	1	Culture Status	0.17	0.97	380	491
	2	Relapse	0.025	0.82	1032	
		Overall	0.025	0.80		

- NB: Conventional approach: max(N) = 1442, $E(N \mid H_0)$ = 348, 68% power

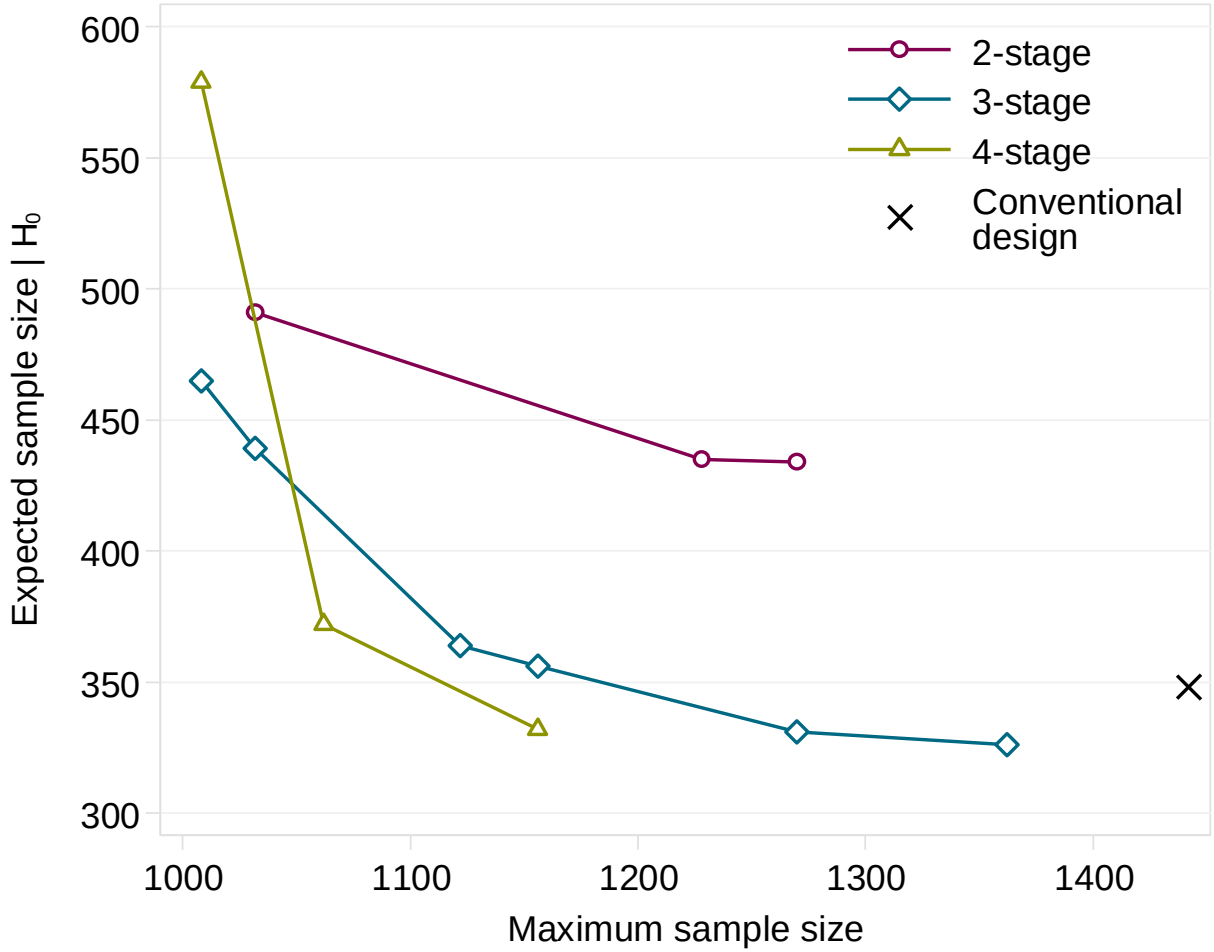
E(N|H_0) vs max(N) of admissible phase II/III TB trial designs



E(N|H_0) vs max(N) of admissible phase II/III TB trial designs



E(N|H_0) vs max(N) of admissible phase II/III TB trial designs



Summary

- There is an increasing need for more efficient trial designs to accelerate drug development

- MAMS designs can achieve this by
    - combining phases 2 & 3 into a single trial
    - testing multiple arms simultaneously
    - ceasing recruitment to poorly performing arms at interim analyses

- We extended the MAMS design for time to event outcomes to allow use of binary intermediate and definitive outcomes

- Further work will allow use of any combination of outcomes (e.g. binary I outcome, continuous D outcome)

Summary

- Admissible designs minimise the weighted sum

$$wE(N|H_0)+(1-w)max(N)$$

  for some $w \epsilon [0,1]$ and are the most ideal feasible designs to use

- They can reduce both expected and maximum sample sizes over the conventional approach of conducting phase 2 & 3 separately, with simultaneous gains in power

- Although effect estimates in multi-stage designs are biased, previous work has shown bias to be negligible if using an intermediate outcome

- Even greater gains in efficiency can be made by evaluating >1 experimental treatment over separate phase 2 & 3 trials for each new therapy

References

- Bratton, D. J., P. P. J. Phillips, et al. (2013). "A multi-arm multi-stage clinical trial design for binary outcomes with application to tuberculosis." Medical Research Methodology **13**:139.
- Dorman, S. E., J. L. Johnson, et al. (2009). "Substitution of moxifloxacin for isoniazid during intensive phase treatment of pulmonary tuberculosis." Am J Respir Crit Care Med **180**(3): 273-280.
- Jung, S. H., T. Lee, et al. (2004). "Admissible two-stage designs for phase II cancer clinical trials." Stat Med **23**(4): 561-569.
- Nunn, A. J., P. P. Phillips, et al. (2008). "Design issues in pivotal drug trials for drug sensitive tuberculosis (TB)." Tuberculosis (Edinb) **88 Suppl 1**: S85-92.
- Phillips, P. P. J., S. H. Gillespie, et al. (2012). "Innovative Trial Designs Are Practical Solutions for Improving the Treatment of Tuberculosis." Journal of Infectious Diseases.
- Royston, P., F. M. Barthel, et al. (2011). "Designs for clinical trials with time-to-event outcomes based on stopping guidelines for lack of benefit." Trials **12**: 81.
- Royston, P., M. K. Parmar, et al. (2003). "Novel designs for multi-arm clinical trials with survival outcomes with an application in ovarian cancer." Stat Med **22**(14): 2239-2256.