A Bayesian Adaptive Design for Clinical Trials in Rare Diseases

Peter Jacko*

joint work with S. Faye Williamson, Thomas Jaki and Sofia S. Villar

IMA and OR Society Conference on
Mathematics of Operational Research

Birmingham, April 20–21, 2017

*Dept. of Management Science, Lancaster University, UK

Clinical Trials

- Two treatments: control (existing one) and novel (not approved yet)
- Is the novel treatment better than the control?
  - ▷ clinically relevant treatment effect difference
  - ▷ if not, it will not be approved!
- One sets up a clinical trial of n patients
- A (clinical trial) design is an allocation policy that specifies which treatment the ith patient will receive

Clinical Trials

- The gold standard design: randomised controlled trial
  - 50% vs 50% fixed equal randomisation
  - avoids all types of biases
  - in use since 1948 (advocated since Hill 1937)
- Its main goal is to learn about treatment effectiveness with a view to prioritising future outside patients
  - maximises power of a treatment effect difference
  - if approved, future patients are, say, 80% confident that the novel treatment is better than the control
- A half of trial patients gets the inferior treatment

Clinical Trials

- Problems with randomised controlled trial
  - ▷ cost: 20% error of not approving a better treatment
    - — development and approval processes: $ billions
  - ▷ faith: once approved, no (simple) way to retract the treatment
    - — worse treatment approved by 5% chance
    - — unforeseen long-term secondary effects
  - ▷ feasibility: requires hundreds of patients for a trial
  - ▷ stationarity: approval process takes years
    - — inappropriate for new diseases and epidemics
  - ▷ ethics: patients join a trial expecting to get a possibly better (unapproved yet) treatment

Multi-Armed Bandit Problem

Multi-armed Bandit Approach

- Maximise healing of patients in the trial
  - ▷ optimally solving exploration/exploitation trade-off
  - ▷ learning takes place during the trial
- The multi-armed bandit motivated by clinical trials
  - ▷ Thompson (Biometrika 1933), Robbins (1952), etc.
- Bandit models are a type of response-adaptive design
- Appropriate model: finite horizon
  - ▷ the celebrated Gittins' theorem does not apply!

Bayesian Bernoulli Bandit Model

- Finite horizon: n sequentially arriving patients
- Two-armed: treatment A or B for each patient
- Binary endpoints: success (1) or failure (0)
- Let X_i and Y_i denote patient i's response from treatment A and B respectively (for $i = 1, ..., n$). Then,

$$X_i \sim \text{Bernoulli}(1,\ \theta_A) \ \text{ and } \ Y_i \sim \text{Bernoulli}(1,\ \theta_B),$$

where θ_A and θ_B are the unknown success probabilities of treatments A and B respectively

Bayesian Approach

- Beliefs $\widehat{\theta}_A$ and $\widehat{\theta}_B$ to be updated over the trial
- Prior Distribution: $\widehat{\theta}_A \sim \text{Beta}(a,\ b)$, $\widehat{\theta}_B \sim \text{Beta}(c,\ d)$ where we take $a = b = c = d = 1$ (uninformative)
- Posterior Distribution: After observing i (j) successes (failures) on treatment A, and k (l) successes (failures) on treatment B, the posterior distribution is represented by another Beta distribution (by conjugacy)

$$\widehat{\theta}_A \sim \text{Beta}(a+i, b+j),\ \widehat{\theta}_B \sim \text{Beta}(c+k, d+l)$$

Optimal Design using DP

- We use dynamic programming (DP) to obtain an optimal adaptive treatment allocation sequence
- Optimal in the sense of maximising the expected total number of successes in the trial
- Specifically, we use backward induction algorithm
- Let $\mathcal{F}_m(i, j, k, l)$ be the expected total number of successes under an optimal policy after m patients
- If $m = n$, there is nothing to do: $\mathcal{F}_n(i, j, k, l) = 0$ $\forall i, j, k, l$

Backward Induction

- If $m = n - 1$ (one patient left):

  1. If treatment A, we compute the expectation

$$\mathcal{F}^A_{n-1}(i,j,k,l) = \frac{i}{i+j} \cdot 1 + \frac{j}{i+j} \cdot 0$$

  2. If treatment B, we compute the expectation

$$\mathcal{F}^B_{n-1}(i,j,k,l) = \frac{k}{k+l} \cdot 1 + \frac{l}{k+l} \cdot 0$$

- We wish to choose the optimal allocation such that

$$\mathcal{F}_{n-1}(i,j,k,l) = \max\{\mathcal{F}^A_{n-1}(i,j,k,l),\ \mathcal{F}^B_{n-1}(i,j,k,l)\}$$

Optimal Design

- Problem? This design is not suitable to implement in practice because it is completely deterministic
- As a result, there is a risk of introducing bias into the trial through the intentional selection of patients (selection bias)
- Therefore, we modify the optimal DP design by forcing actions to be randomised
  - ▷ see also Cheng & Berry (Biometrika, 2007)
- Helps to maintain blinding and reduce the risk of bias

Forcing Randomised Actions

- Action 1: treatment A is allocated with probability p
- Action 2: treatment B is allocated with probability p
- The expected total number of successes under Action 1

$$\mathcal{V}_m^1(i,j,k,l) = p \cdot \mathcal{F}_m^A(i,j,k,l) + (1-p) \cdot \mathcal{F}_m^B(i,j,k,l)$$

- The objective function becomes

$$\mathcal{V}_m(i,j,k,l) = \max\left\{\mathcal{V}_m^1(i,j,k,l),\ \mathcal{V}_m^2(i,j,k,l)\right\}$$

- Lower selection bias, but lower controllability

Randomised Variant

- Problems? After running simulations, we found:
  - ▷ this design is very underpowered for high p
  - ▷ in some of the runs (only a few out of 10,000), all patients were allocated to only one of the treatments
- This means we cannot be confident about the results
- ...we cannot calculate important performance measures
- Therefore, we lower-limit the number of observations on each treatment

Constrained Variant

- We modify the optimal randomised DP policy by adding a constraint to ensure that we obtain $\geq \ell$ observations from each treatment
- To do this, we assign a large penalty to every terminal state that has $< \ell$ observations on a treatment arm
- The undesirable states will now be avoided (as much as possible) by the optimal policy
- We tried a range of values for ℓ, i.e. $0.05n$, $0.10n$, $0.15n$, $0.20n$ and $0.25n$. (Note that $0.50n$ corresponds to equal, fixed randomisation)

Simulation Study

- We evaluate the performance of proposed designs by
  - Bias of the treatment effect estimator
  - ...and its mean squared error (MSE)
  - Statistical power
  - Expected proportion of successes (EPS)
  - Patients allocated to the superior arm (On sup)
- For each configuration, we replicate 10,000 trials

Simulation Results: Randomised Variant

Example. $n = 75, \theta_A = 0.2, \theta_B = 0.8$

p	Bias	MSE	Power	EPS	On sup
50%	0.001	0.004	1.000	50.0%	50.0%
60%	0.001	0.005	1.000	55.7%	59.6%
70%	0.001	0.007	0.999	61.5%	69.2%
80%	0.004	0.010	0.995	67.2%	78.8%
90%	0.009	0.019	0.937	73.0%	88.3%
100%	0.100	0.043	0.118	78.6%	97.6%

- The Power (almost) does not change if p increased from 50% to 60% or 70%. Room for increasing EPS!

Simulation Results: Randomised Variant

Example. $n = 75, \theta_A = 0.2, \theta_B = 0.6$

p	Bias	MSE	Power	EPS	On sup
50%	0.001	0.004	0.938	40.0%	50.0%
60%	0.002	0.005	0.935	43.7%	59.1%
70%	0.002	0.007	0.910	47.3%	68.2%
80%	0.005	0.009	0.830	50.9%	77.3%
90%	0.015	0.015	0.636	54.4%	86.0%
100%	0.089	0.030	0.070	57.7%	94.2%

- The Power (almost) does not change if p increased from 50% to 60% or 70%. Room for increasing EPS!

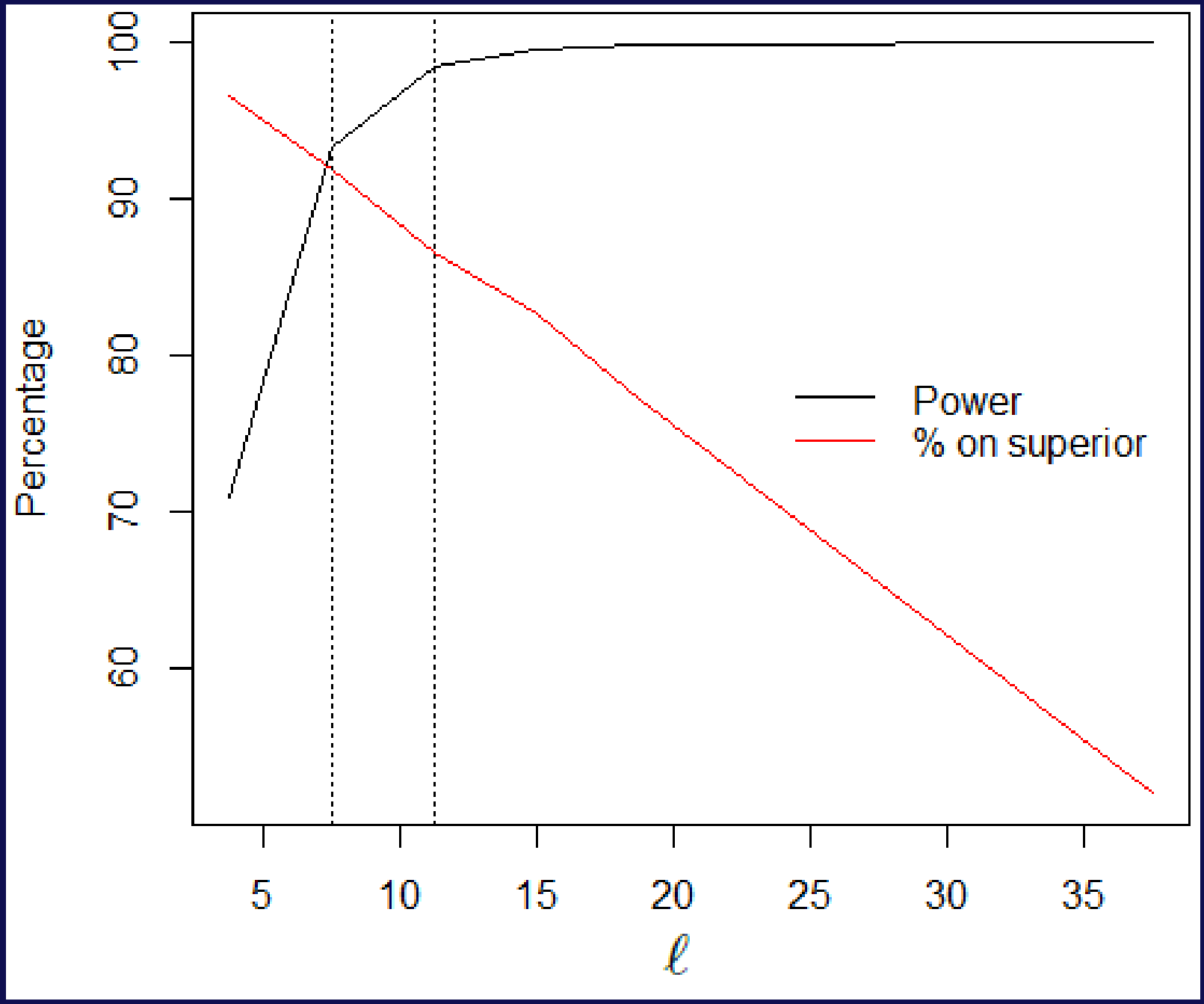
Simulation Results: Constrained Variant

Example. $n = 75, \theta_A = 0.2, \theta_B = 0.8$

ℓ	Power	EPS	On sup
$0.05n$	0.442	78.0%	96.6%
$0.10n$	0.884	75.2%	91.9%
$0.15n$	0.964	72.1%	86.7%
$0.20n$	0.985	69.7%	82.7%
$0.25n$	0.997	66.5%	77.3%
$0.50n$	1.000	51.2%	52.0%

- As ℓ increases, the power of the design increases hyperbolically, but the EPS and % allocated to the superior arm decreases linearly

Simulation Results: Constrained Variant



Simulation Results: Constrained Randomised Variant

Example. $n = 75, \ell = 0.15n, \theta_A = 0.2, \theta_B = 0.8$

p	Bias	MSE	Power	EPS	On sup
60%	0.001	0.005	1.000	55.7%	59.6%
70%	0.001	0.007	0.999	61.5%	69.2%
80%	0.003	0.010	0.996	67.2%	78.7%
90%	0.003	0.014	0.977	71.3%	85.5%

- The Power is quite high even if p increased to 80% or 90%. Also bias diminishes!

Simulation Results: Constrained Randomised Variant

Example. $n = 75, \ell = 0.15n, \theta_A = 0.2, \theta_B = 0.6$

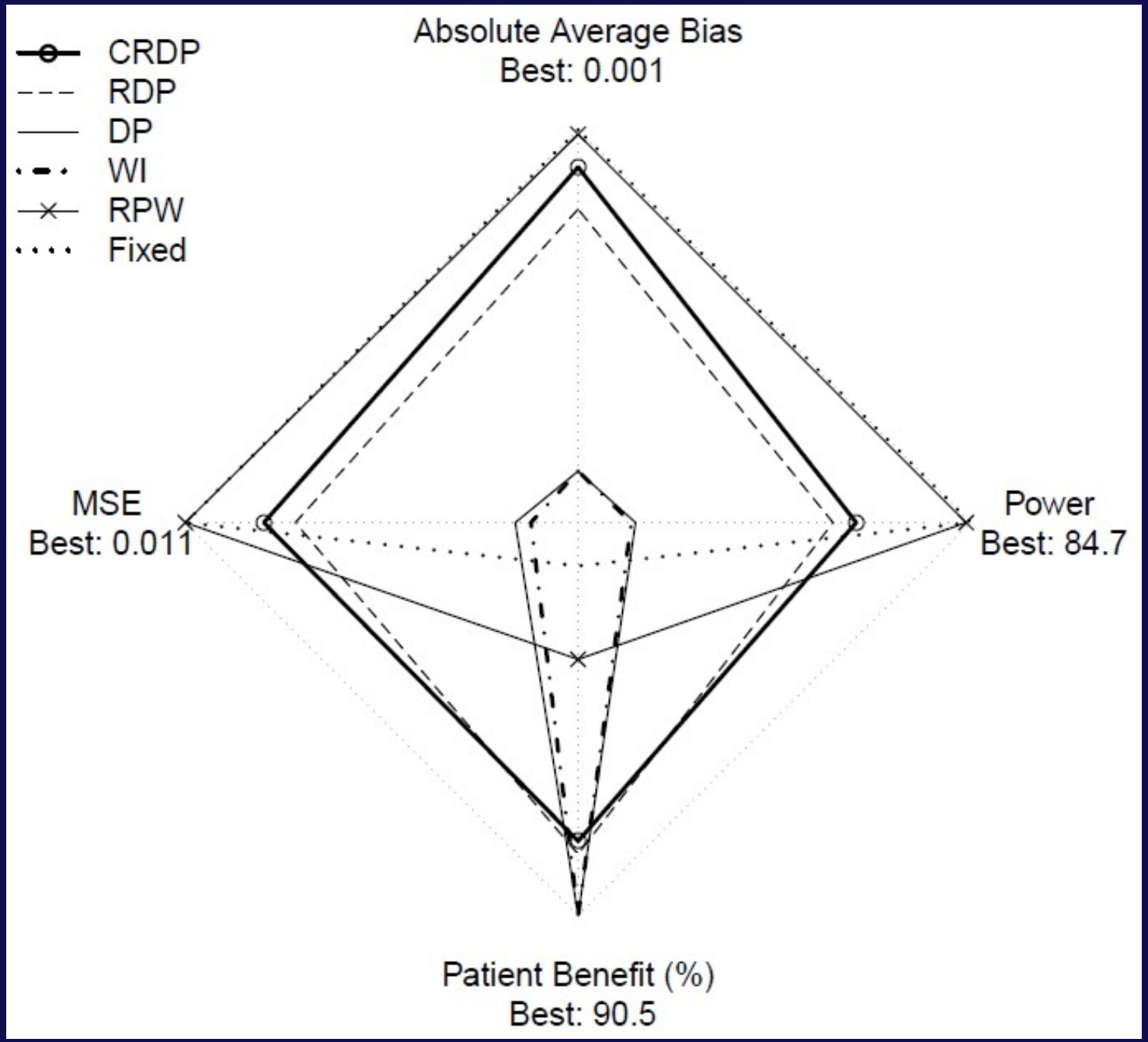
p	Bias	MSE	Power	EPS	On sup
60%	0.002	0.005	0.935	43.7%	59.1%
70%	0.002	0.007	0.910	47.3%	68.2%
80%	0.005	0.009	0.834	50.9%	77.2%
90%	0.008	0.013	0.724	53.6%	84.0%

- The Power is quite high even if p increased to 80% or 90%. Also bias diminishes!

Simulation Study

- We compare our proposed constrained randomised variant of DP (CRDP) design to the following designs:
  - ▷ Fixed randomisation (the gold standard)
  - ▷ Randomised play-the-winner rule (RPW)
  - ▷ Optimal dynamic programming policy (DP)
  - ▷ Whittle index policy (WI)
  - ▷ Randomised variant of the DP policy (RDP)
- We suggest to set $p = 90\%, \ell = 0.15n$ in CRDP

Simulation Results: Designs Comparison



Simulation Results: Designs Comparison

Example. $n = 75, \theta_A = 0.2, \theta_B = 0.8$

Design	Bias	Power	EPS	On sup
Fixed	0.000	1.000	50.0%	50.0%
RPW	0.008	0.998	66.2%	76.9%
WI	0.098	0.108	78.6%	97.6%
DP	0.100	0.118	78.6%	97.5%
RDP	0.009	0.937	73.0%	88.3%
CRDP	0.003	0.977	71.3%	85.5%

Simulation Results: Designs Comparison

Example. $n = 75, \theta_A = 0.2, \theta_B = 0.6$

Design	Bias	Power	EPS	On sup
Fixed	0.000	0.935	40.0%	50.0%
RPW	0.002	0.928	46.2%	65.4%
WI	0.092	0.066	57.8%	94.4%
DP	0.088	0.074	57.7%	94.1%
RDP	0.015	0.636	54.4%	86.0%
CRDP	0.008	0.724	53.6%	84.0%

Simulation Results: Designs Comparison

- In our suggested CRDP design
  - the % expected proportion of successes is much higher than in the traditional fixed and RPW designs
  - the % allocated to the superior arm is much higher than in the traditional fixed and RPW designs
  - the power is largely improved upon relative to the other bandit designs
  - the bias is negligible, opposed to large bias of other bandit designs

Conclusion

- We address some of the key issues preventing bandit models from being implemented in clinical trial practice
  - ▷ lack of randomisation
  - ▷ insufficient statistical power
  - ▷ biased estimates of the treatment effect
- We need to talk to statisticians and clinicians about bandit models
  - ▷ give me randomisation probability and desired power
  - ▷ I tell how to randomise treatments to heal patients

Thank you for your attention

...and see you in Lancaster

- The 7th meeting of the EURO WG on Stochastic Modelling
- 13–15 June 2018, Lancaster University
- Become member at www.stochmod.eu — it's free!