Clinical Trial Design for Rare Diseases using Bayesian Bandit Models

Faye Williamson [1], Peter Jacko [1], Sofia Villar [2], Thomas Jaki [1]

[1]STOR-i CDT, Lancaster University, [2]MRC Biostatistics Unit, Cambridge.

Introduction & Motivation

- Consider a two-arm clinical trial with binary end points and a finite number of patients, n.
- Suppose each treatment, A and B, has an unknown success probability, θ_A and θ_B, respectively, and each patient's response is immediately available.
- The null and alternative hypotheses are formulated as

$$H_0 : \theta_A = \theta_B \text{ versus } H_1 : \theta_A \neq \theta_B,$$

and we assume $\theta_A,\ \theta_B \sim \text{Beta}(1,\ 1)$ a priori.

Objective: To design a clinical trial which identifies the superior treatment (explores) whilst effectively treating the trial participants (exploits). This will be particularly useful in trials for rare diseases.

- This is a natural application area for bandit models which seek to balance the exploration versus exploitation trade-off to obtain an optimal allocation policy.
- Bandit models are a type of response-adaptive design.
- Learning takes place during the trial (rather than just at the end as in the traditional randomised controlled trial).
- The optimality property, in terms of maximising the expected number of patient successes, is the primary motivation behind implementing bandit-based designs in clinical practice.

Optimal Design using Dynamic Programming (DP)

- We use DP to obtain the optimal adaptive treatment allocation sequence.
- The idea behind DP is a recurrence equation (the Bellman equation), which relates the expected total reward at a given decision time to the distribution of its possible values at the next decision time.
- We implement a backward induction algorithm in which we start with patient n and proceed iteratively towards the first patient.

Limitations for Trial Design

- This design is completely deterministic.
- Optimal designs which achieve the highest patient benefit suffer from the lowest power.

We focus on modifications to the optimal design which aim at overcoming these limitations without having a significant impact on the patient benefit criterion.

Randomised Dynamic Programming Design (RDP)

- A natural first step is to modify the optimal DP design by forcing actions to be randomised.
- This helps to maintain blinding and reduce the risk of bias.
- We define the following actions so that each treatment has a probability of at least $1 - p$ of being allocated to each patient, where $0.5 \leq p \leq 1$.

1. **Action 1:** A patient receives treatment A with probability p (and treatment B with probability $1 - p$).
2. **Action 2:** A patient receives treatment B with probability p (and treatment A with probability $1 - p$).

- We tried a range of values for p and suggest setting $p = 0.90$.
- This design markedly improves power and trades a small reduction in optimality for randomisation.

Further Limitations

- There is a possibility that all patients may be allocated to only one treatment.

Constrained RDP Design (CRDP)

- We propose a constrained variant of the RDP design which ensures that we always obtain at least Y observations from each treatment arm.
- Therefore, we can no longer end up with no observations on a treatment arm.
- To do this, we assign a large negative penalty to every terminal state that has less than Y observations on a treatment arm.
- This causes the undesirable states to now be avoided.
- We tried several values for the lower bound Y and suggest setting $Y = 0.15n$.

Simulation Study

We evaluate the CRDP design in several scenarios by simulating 10,000 replications and focusing on the following performance measures:

- Power; type I error rate; average bias of the treatment effect estimator; mean squared error (MSE) and the percentage of patients allocated to the superior treatment (% on sup).

...and we compare our proposed CRDP design to the following designs:

- Traditional fixed randomisation; randomised play-the-winner (RPW); Whittle index policy (WI), DP and RDP.

These figures correspond to $n = 75$, $\theta_A = 0.5$ and $\theta_B \in (0.1,\ 0.9)$.

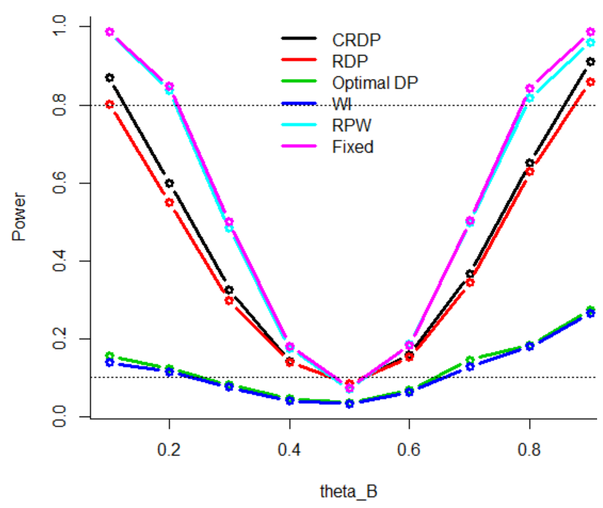

Figure 1: Power and type I error

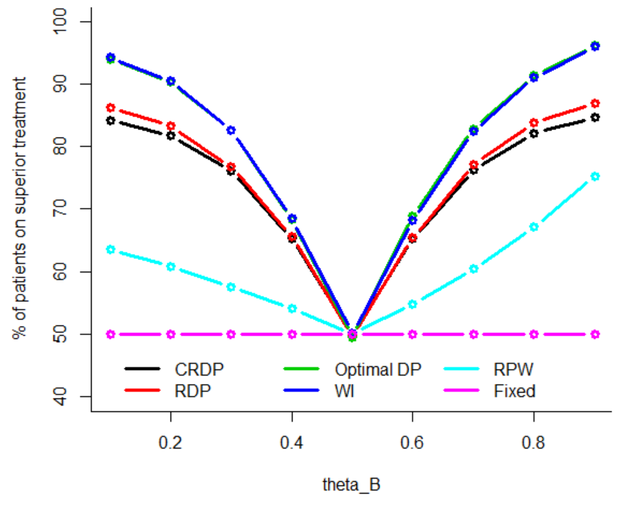

Figure 2: % on sup

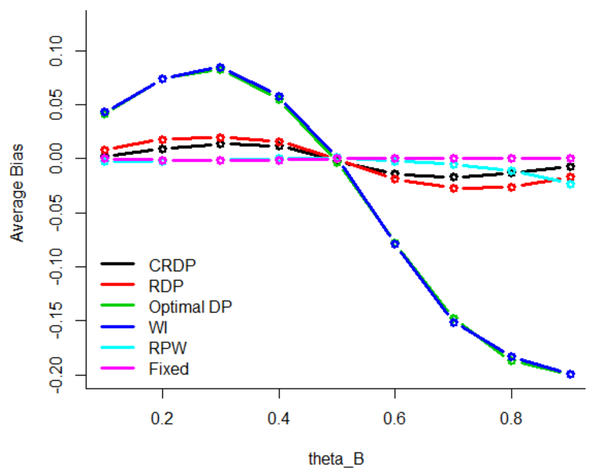

Figure 3: Average bias

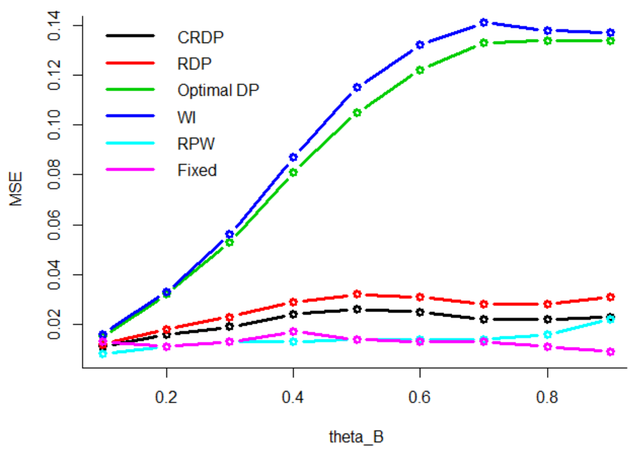

Figure 4: MSE

Conclusions

Our proposed CRDP design produces very promising results:

1. The power is greatly improved upon relative to the other bandit designs (Figure 1).
2. The % of patients allocated to the superior treatment is much higher than in the traditional fixed and RPW designs (Figure 2).
3. The bias and MSE of the treatment effect estimator is greatly reduced compared to the other bandit designs (Figures 3 and 4).

Such designs will be particularly useful for rare diseases in which a substantial proportion of patients exhibiting the disease are included in the trial, and therefore the priority is to treat these patients as effectively as possible.

http://www.lancaster.ac.uk/pg/williasf/ s.f.williamson@lancaster.ac.uk