

Evaluating a Responsive Adaptive Clinical Trial using simulations

Moaz Sidat¹. Supervisor: Holly Jackson²

¹Lancaster University, ²STOR-i

1 Background

- The development of medical treatments require extensive research and testing to determine effectiveness, safety and the ability of the treatment to work.
- Clinical trials are the testing of treatments on humans. **Randomised Controlled Trials (RCTs)** are the most often used approach. RCTs assign each patient to a treatment with each treatment having an equal probability.
- Although RCTs have a good ability to identify a significant difference between the treatments, the number of patients successfully treated in the trial is not particularly high.

2 Introduction

- The proposed idea is called the **Response Adaptive Design (RAD)** which uses information from the previously treated patients to decide the allocation of the next patient. This is an example of the **Multi Armed Bandit problem** in the clinical trials context. The main objective is to balance the number of patients treated successfully (**earning**) and identifying the best treatment (**learning**).
- An issue with RADs is that characteristics such as weight, blood type, age (referred to as **covariates**), could alter the outcome of treatments.
- In this study, we will be comparing the RCT and RAD designs in their ability to detect a difference in treatments and the number of patients successfully treated.

3 Method

Suppose we have N patients in the trial. The method used in this simulation study is:

1. Allocate the first n patients equally to each treatment and record the outcome.
2. Estimate the predicted outcome of patient $n + 1$ using a regression method, the **Nearest Neighbour Method (NNM)** based on their covariate.
3. Assign this patient to the treatment with the best predicted outcome, with high probability.
4. Given we know the outcome of patient $n + 1$ update our information and repeat steps **2** and **3** for the next patients.

4 Power and Proportion

- **Power**- Consider the hypotheses:

- $H_0: \theta_1 = \theta_2$. (There is no difference in treatments).
- $H_1: \theta_1 \neq \theta_2$. (There is a difference in the treatment).

Where θ_1 and θ_2 are the effects of the two treatments. Hence the power is the probability of detecting an effect in the treatment where there is one. The power will be used as a measure of the ability of the method to detect a significant difference in treatments.

- **Proportion**- The proportion of patients on their best treatment will be used as a measure of patients successfully treated.
- The desired outcome is the design having a **high proportion and high power**.

6 Conclusion

- RAD method predominantly has a higher proportion of patients on their best treatment than the RCT method.
- The values of power varies with the scenario. Majority of the time, RCTs have a higher power than the RAD method.
- The large difference in power of Scenario 4 maybe due to simulation error since we have only conducted 100 simulations.
- It could be possible that the type of outcome affects the ability of the trial to detect a difference in treatments.

7 Further Work

- Outcomes of treatments are not known immediately which this method assumes.
- Try different regression techniques within the RAD method.
- In the case of rare diseases, the patients in the trial are a higher proportion of the total population suffering the disease so we would want as many patients treated as we can.

8 References

[Yang and Zhu, 2002] Yang, Y. and Zhu, D. (2002). Randomized allocation with nonparametric estimation for a multi-armed bandit problem with covariates. *Ann. Statist.*, 30(1):100–121.

Contact Information

- Email: m.sidat@lancaster.ac.uk

5 Simulation Study and results

In our simulation study, we have 100 patients with 2 treatments and have simulated 100 times.

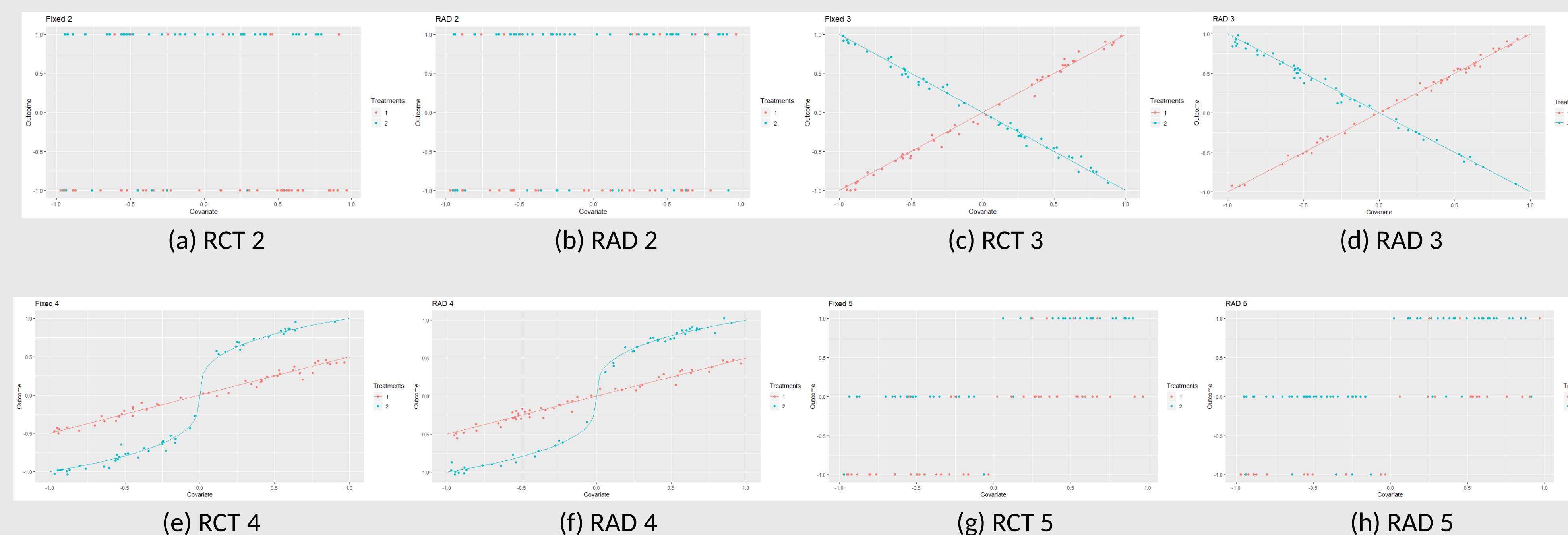
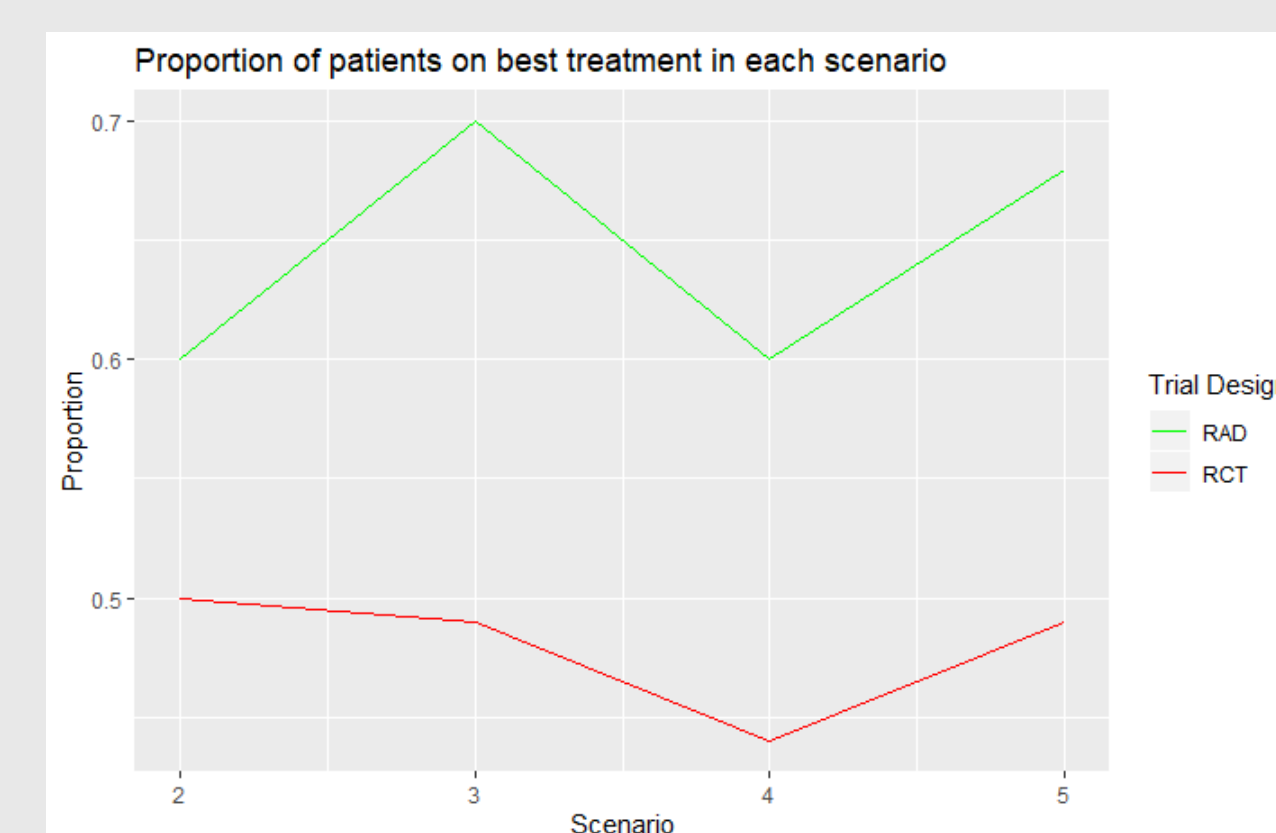
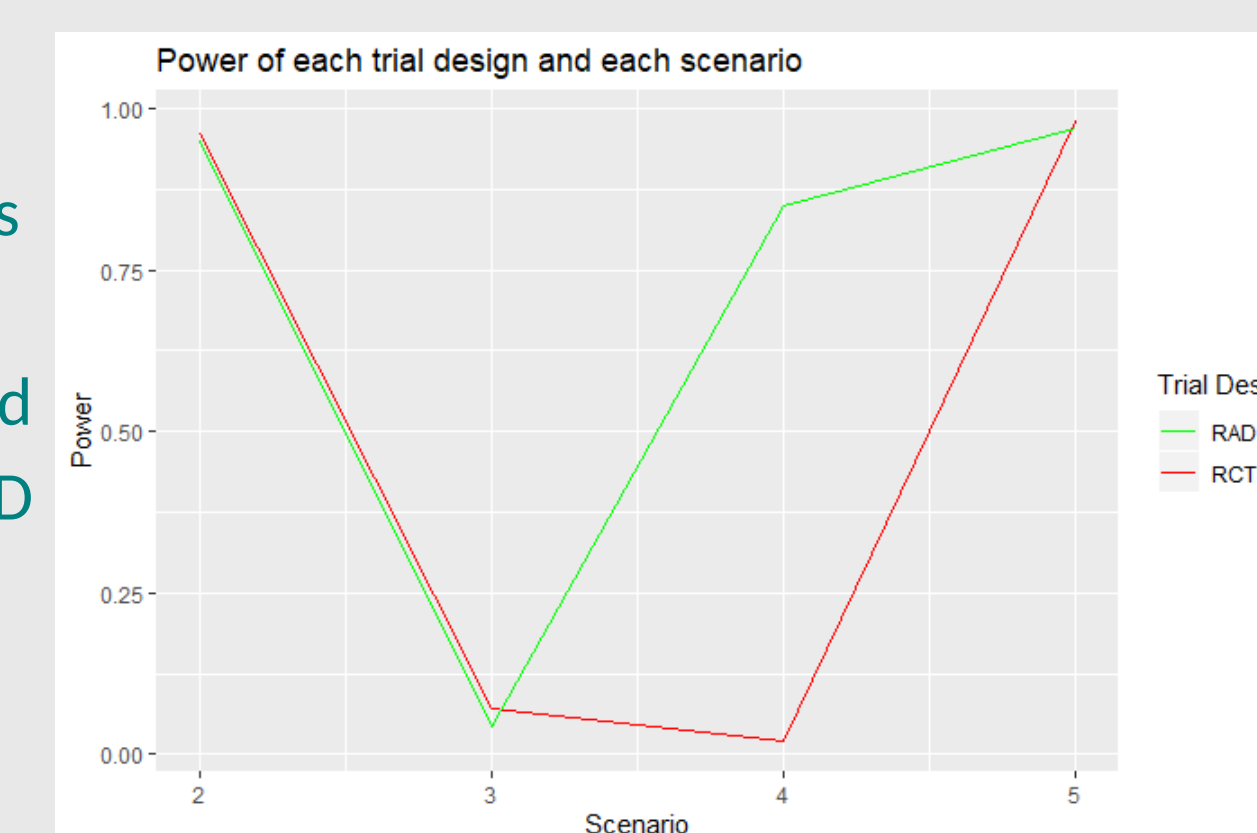


Figure: (a)(b) represent a scenario with a **binary outcome**, (c)(d), (e)(f) represent scenarios with a **continuous outcome** and (g)(h) represents the scenario with a **categorical outcome**.



The average proportions for the scenarios investigated are 0.48 and 0.65 for the RCT and RAD methods respectively.



The power seems to vary depending on the scenario. Beside scenarios 4, the RCT method has a higher power but not by a huge margin.