ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Economic Evaluation

# A Tutorial on Value-Based Adaptive Designs: Could a Value-Based Sequential 2-Arm Design Have Created More Health Economic Value for the Big CACTUS Trial?

Laura Flight, PhD, Alan Brennan, PhD, Isabelle Wilson, MSc, Stephen E. Chick, PhD

## ABSTRACT

*Objectives:* Value-based trials aim to maximize the expected net benefit by balancing technology adoption decisions and clinical trial costs. Adaptive trials offer additional efficiency. This article provides guidance on determining whether a value-based sequential design is the best option for an adaptive 2-arm trial, illustrated through a case study.

*Methods:* We outlined 4 steps for the value-based sequential approach. The case study re-evaluates the Big CACTUS trial design using pilot trial data and a model-based health economic analysis. Expected net benefit is computed for (1) original fixed design, (2) value-based design with fixed sample size, and (3) optimal value-based sequential design with adaptive stopping. We compare pretrial modeling with the actual Big CACTUS trial results.

*Results:* Over 10 years, the adoption decision would affect approximately 215 378 patients. Pretrial modeling shows that the expected net benefit minus costs are (1) £102 million for the original fixed design, (2) £107 million (+5.3% higher) for the value-based design with optimal fixed sample size, and (3) £109 million (+6.7% higher) for the optimal value-based sequential design with maximum sample size of 435 per arm. Post hoc analysis using actual Big CACTUS trial data indicates that the value-adaptive trial with a maximum sample size of 95 participant pairs would not have stopped early. Bootstrap simulations reveal a 9.76% probability of early completion with n = 95 pairs compared with 31.50% with n = 435 pairs.

*Conclusions:* The 4-step approach to value-based sequential 2-arm design with adaptive stopping was successfully implemented. Further application of value-based adaptive approaches could be useful to assess the efficiency of alternative study designs.

*Keywords:* adaptive clinical trial, Bayesian trial design, expected value of sample information, sequential clinical trial, value-based trial design.

VALUE HEALTH. 2024; 27(10):1328–1337

## Highlights

- Value-based trials seek to balance the cost of a clinical trial with the expected health and economic benefits of the subsequent technology adoption decisions. Applying this approach to the adaptive design context has the potential to better balance the costs of clinical trials with the health benefits of health technology adoption decisions that are based on the results of those trials.
- We bridge the gap between theory and practice by providing step-by-step guidance to implementing a value-based sequential design for an adaptive 2-arm trial. We demonstrate the first application of the value-based sequential design that uses data from an external pilot trial and do so in the context of publicly funded research.
- With this tutorial, we hope to increase experience and application of these approaches so that they can be implemented with only small changes to current skills and practices. We also demonstrate mechanisms to assess the health economic value of doing a value-adaptive sequential trial compared with some other clinical trial designs.

## Introduction

Value-based trials seek to balance the cost of a clinical trial with the expected health and economic benefits of technology adoption decisions made with trial evidence,[1] making them appealing to healthcare providers and funders such as the United Kingdom (UK) National Health Service.[2] Value-adaptive designs (see Flight et al[3]) extend this approach to adaptive trials. They explicitly consider the expected benefits and costs of adapting a trial to maximize expected population health for the research money spent.[4]

A proposed approach is the value-based sequential 2-arm design with adaptive stopping.[4,5] This design calculates the expected health and economic value of gathering further information about the cost-effectiveness of 1 health technology compared with another, as the trial progresses, informing stop-go decisions.[4] It has been applied in 2 retrospective case studies using the observed trial data[6,7] (chapter 3).

This article bridges the gap between theory and practice by providing step-by-step guidance for determining whether a value-based sequential design is the best option for an adaptive 2-arm trial. We operationalize this approach in a (UK) case study, using the National Institute for Health Research (NIHR)–funded "Big CACTUS" trial. This provides the first application of a value-based sequential design using pilot trial data and extends existing applications of the approach from a within-trial economic analysis to a model-based health economic analysis.[6,7]

## Methods

### Overview of Value-Based Sequential 2-Arm Designs With Adaptive Stopping

Using Bayesian decision theory, Chick et al[4] define a value-based sequential design for a 2-arm trial with adaptive stopping. Participants are randomized pairwise and sequentially to new technology or control. Pretrial beliefs about effects and costs are used to provide a previous distribution for the expected incremental net monetary benefit (INMB) of the new technology versus a control. The objective is to maximize the expected INMB of the technology adoption decision minus the cost of the trial.

When the expected benefits of continuing to randomize participants are not deemed worth the expected costs the trial stops. Given that data on INMB arrive from participants reaching the health economic follow-up time point, the posterior mean for the expected INMB is calculated. If this lies in the continuation region, within the upper and lower stopping boundary, the trial continues. Crossing the stopping boundary or reaching the maximum sample size halts randomization, and enrolled participants are followed-up.

### Four Steps to Apply the Value-Based Sequential Approach

Four steps to implement the value-based sequential approach are detailed below (Box 1).

**BOX 1.** Summary of steps to compare multiple trial designs including a value-based sequential design for a 2-arm trial with adaptive stopping.

1. Specify the key parameters required for the trial designs under consideration.
   a. Willingness to pay threshold
   b. Fixed and variable costs of conducting the research and adopting the technology into practice
   c. Size of the population to benefit from the health technology adoption decision
   d. Sampling variance of the INMB in the population of participants considered
   e. Delay between randomization and observing the measure of INMB for each pair of participants
   f. Previous mean and variance of the expected INMB
2. Compute the optimal trial design within each of the design structures under consideration.
   a. Traditional fixed sample size design
   b. Value-based fixed sample size design
   c. Value-based sequential design
3. Compare the candidate trial designs and select a design for the trial.
   a. Compare the expected value and net benefit of competing designs.
   b. Consider practical issues such as being able to recruit the required population, staffing concerns, and financial factors.
   c. Use simulations to explore the characteristics such as expected sample size, expected cost, and frequentist type I and type II error rates.
4. Conduct the trial with the chosen design.

Additional considerations and resources may be required if a sequential design is adopted such as additional interim analyses and greater health economic involvement at the design stage of the trial.

INMB indicates incremental net monetary benefit.

#### Step 1: Specify key parameters required for the trial designs under consideration

We define 6 key parameters, informed by the research proposal, pretrial data (eg, from a pilot trial), discussions with clinical researchers, and following existing guidance for good practice in technology appraisals.[8]

a. Willingness to pay threshold—the amount a decision maker is willing to pay for one unit of effectiveness. We used £20 000 for one quality-adjusted life-year (QALY) based on UK guidance.[8]
b. Fixed ($C_f$) and variable costs ($C_v$) of conducting the research. Fixed costs include any costs incurred in the setup, conduct, and analysis and after the end of the trial regardless of the trial design and sample size. This can include things such as dissemination costs. These costs are also independent of whether the technology is adopted or not. We have assumed the trial results will be reported and disseminated whether the trial suggests the new technology is cost-effective or not. Variable costs include costs associated with randomizing, monitoring, and delivering the intervention to participants.[9]
c. Size of the population that will benefit from the health technology adoption decision ($P$). This can be estimated by multiplying the estimated annual incidence of the clinical condition by the time horizon over which the adoption decision is deemed to apply,[4] incorporating discounting if appropriate.[10] For example, in some jurisdictions the QALYs gained 10 years from now are considered to be less valuable than QALYs gained immediately, and so discounting is applied to the population that will be benefiting in future years.
d. Sampling variance ($\sigma_X$) of the INMB in the population of participants considered based on pilot data and other background information
e. Delay ($\tau$), measured in participant pairs, between randomization and observing the measure of INMB for each participant—calculated by multiplying the expected recruitment rate by the time to follow-up for the cost-effectiveness data
f. Previous mean ($\mu_0$) and SD ($\sigma_0$) of the expected INMB based on pilot data and other background information

Estimating ($\mu_0$), ($\sigma_0$), and ($\sigma_X$) may require the setup of evidence synthesis on previous existing data and knowledge on the technology and comparator together with a health economic model incorporating uncertainty. Estimates for $\sigma_X$ may be updated as evidence accumulates.[11]

#### Step 2: Compute the optimal trial design within each of the design structures under consideration

Traditional frequentist power calculation methods can be used to determine a fixed sample size for a traditional design.[12,13] Although this step is not needed to design a value-adaptive trial, we do so here for comparison.

A fixed sample size design incorporating QALY and cost information can also be computed using expected value of sample information (EVSI) calculations to determine the expected value of different proposed sample sizes using well-established methods[14-17] and approaches to find the optimal fixed sample size, for example, https://github.com/andres-alban/ValueBasedTrials.[18] The same 6 key parameters in Section 2.2.1 can be used to produce the expected net benefit of sampling (ENBS) for each proposed sample size by calculating the population EVSI minus the estimated cost of carrying out the research.[19] The sample size with the highest ENBS can be recommended as the most cost-effective fixed sample size design.[1,10]

An adaptive design can be implemented with stopping boundaries computed using the optimal value-based sequential model as outlined in Section 2.1. This calculates the EVSI and ENBS in a way that accounts for adaptive stopping times based on the theory of dynamic programming as in Chick et al[4] and computed using associated MATLAB code.[20] In their notation, the expected value of making a technology adoption decision now (before conducting a trial) using current information is $\mu_0 \times P$ if the previous INMB for the new treatment is positive. The expected value of collecting further information using the value-based sequential design is denoted $V^*(\mu_0)$. Thus, the ENBS for the optimal value-based sequential design is $V^*(\mu_0) - \max(0, \mu_0 \times P)$. After the trial starts, a similar computation is made to see whether the trial should continue or randomization stopped.

#### Step 3: Compare the candidate trial designs and select a design for the trial

Once a range of designs have been calculated, the research team can make an informed decision on the most appropriate design. We consider 3 options:

a) traditional fixed sample size design
b) value-based design with a fixed sample size
c) value-based sequential design

and compare each against not conducting a trial and selecting the technology based on previous existing knowledge.

The value-based approach highlights the opportunity to select a design that maximizes the ENBS. In practice, the final judgment on implementation will consider additional issues such as successful recruitment, staffing concerns, financial factors, and public and participant involvement.[21]

Opting not to conduct further research incurs no additional costs, whereas proceeding with research, regardless of the chosen trial design, involves fixed costs (Section 2.2.1). Variable costs per participant are incurred once the trial starts. Although our case study assumed equal variable costs per participant pair, these costs might vary based on the trial design.

Additional performance metrics for the trial designs can also be computed alongside EVSI and ENBS. Bootstrapping or parametric simulations can be used to explore the characteristics of the candidate trial designs including sample size/expected sample size, expected cost of the study, and probability that the new technology is cost-effective. It is also possible to produce frequentist power curves for the case of value-based multiarm trials.[22]

#### Step 4: Conduct the trial with the chosen design

Upon choosing a design, the trial proceeds, and outputs are analyzed. For a value-based sequential design, accruing data are used to calculate incremental costs and incremental effects for the participant pair and update the estimate of posterior expected INMB to determine potential early trial termination. This is based on the observed data collected so far in the trial. Compared with a fixed design, this requires additional data extraction, cleaning, analysis, and reporting.[23] Further practical considerations are discussed by Forster et al[6,7] and Flight et al.[3,9]

## Case Study

### Overview of Approach

We follow 4 steps in a case study based on a real trial with a fixed sample size design. Pretrial analysis (steps 1-3) uses only pretrial data to explore the potential benefits of a value-adaptive design. We illustrate step 4 using the trial's actual data and insights are gained through bootstrapping simulations on the accrued data for the value-adaptive designs.

### The CACTUS Pilot Trial, Big CACTUS Trial, and Cost-Effectiveness Model

The case study concerns stroke survivors with aphasia, examining a new intervention—computerized speech and language therapy (CSLT). The intervention was piloted in the 2-arm CACTUS pilot trial of CSLT versus usual care (UC).[24] The subsequent Big CACTUS trial was a traditional fixed sample size study that recruited 278 participants to 3 treatment arms—CSLT, UC, and attention control.[25] Big CACTUS was funded by the NIHR Health Technology Assessment Programme.[26] The results found a significant difference in effectiveness between CSLT and UC. Long-term cost-effectiveness of CSLT was assessed using an Excel[27] model-based cost-utility analysis (see Chapter 5[28]).

### Case Study Options for Trial Designs

Our case study re-examined the design of Big CACTUS, imagining that we have decided to run a 2-arm trial of CSLT versus UC and we want to consider different trial design options. We used the data from the CACTUS pilot trial to specify the previous distributions, and together with a health economic model, we then apply steps 1 to 3 to calculate the a priori expected performance of 4 designs:

1. original trial (traditional, fixed sample size) with $n_a$ pairwise allocations
2. value-based fixed sample size design that maximizes ENBS with a sample size of $n_b$ pairwise allocations
3. 2 value-based sequential designs with the maximum number of pairwise allocations taking 2 values:
    - $n_c = n_a$ (from the original trial),
    - $n_c = n_b$ (from the optimal value-based fixed sample size design).

In practice, one may consider value-based sequential designs with any maximum sample size that is practicable given recruitment rates and project duration constraints.

Full methods are below and in Supplemental Material (Section 6) found at https://doi.org/10.1016/j.jval.2024.06.004. The code will be made available on the Economics of Adaptive Clinical Trials website (https://www.sheffield.ac.uk/ctru/completed-trials/enact#Software).

### Replication of Cost-Effectiveness Model in R

For simplicity, our reanalysis of the Big CACTUS trial design focuses on the CSLT and UC arms only, reflecting the new intervention compared with standard practice in the UK. The analysis uses the multiple imputation data set from the original trial.[28] The original Excel-based model was replicated in R[29] for integration with other code (see Latimer et al[30] and its supplementary material). Briefly, the model is a standard Markov cohort model with a cycle length of 3 months, and health states "Aphasia," "Good response," and "Death." "Good response" is defined by a clinically meaningful 10% improvement from baseline in either of the

**Table 1.** Parameter values for the CACTUS case study, restricted to 2 arms, CSLT and UC.

| Definition | Value | Source |
|---|---|---|
| Step 1a: Willingness to pay threshold | | |
| Maximum willingness to pay for one QALY | £20 000 | National Institute for Health and Care Excellence[8] |
| Step 1b: Fixed and variable costs of conducting the research | | |
| Fixed cost of adopting CSLT | £0 | Assumption |
| Estimated spend on fixed costs before starting trial | £581 765 | Big CACTUS grant application |
| Estimated spend during recruitment (per participant) | £1647 | Big CACTUS grant application |
| Estimated spend during follow-up (per participant) | £706 | Big CACTUS grant application |
| Estimated spend on fixed costs post follow-up | £193 044 | Big CACTUS grant application |
| Total spend on fixed costs | $C_f$ = £774 809 | Big CACTUS grant application |
| Total spend | $TC$ = £1 445 565 | Big CACTUS grant application for 3-arm trial |
| Estimate variable cost per allocation to 2 arms | $C_v$ = £4706 | Big CACTUS grant application |
| Fixed cost of adopting CSLT into practice | £0 | Assumption |
| Step 1c: Size of the population that will benefit from the health technology adoption decision | | |
| Population expected to benefit from adoption decision (over a 10-year time horizon) | $P$ = 215 378 | Big CACTUS HTA report[28] |
| Step 1d: Sampling variance of the INMB in the population of participants considered | | |
| SD for INMB in population | $\sigma_X$ = £10 895 | Supplementary material |
| Step 1e: Delay between randomization and observing the measure of INMB for each participant | | |
| Delay for observing INMB endpoint (years) | 1 | Big CACTUS HTA report[28] |
| Delay for observing INMB endpoint (participant pairs) | $\tau$ = 55 | Annual rate of recruitment |
| Step 1f: Previous mean and variance of the expected INMB | | |
| Effective sample size of the previous distribution | $n_0 = 7$ | Pairwise allocations in CACTUS pilot[24] |
| Previous mean for unknown expected INMB (EINMB) | $\mu_0$ = £3190 | Our implementation of the CACTUS project's cost-effectiveness model,[31] and CACTUS pilot data[24] |
| Previous SD for unknown EINMB | $\sigma_0$ = £4118 | Supplementary material |

CSLT indicates computerized speech and language therapy; HTA, health technology assessment; INMB, incremental net monetary benefit; QALY, quality-adjusted life-year; UC, usual care.

coprimary outcome measures (number of words found correctly of 100 personally relevant words; increase of ≥0.5 points on Therapy Outcome Measures activity scale). Costs for the intervention were based on staff activity logs from the trial. UC costs were omitted because both arms include UC. Utilities for each health state were based on analysis of individual participant's EQ5D at 6, 9, and 12 months. The model outputs are discounted mean costs and mean QALYs.

## *Case Study Step 1: Specifying the Key Parameters Required for the Designs*

Parameter values required to implement the value-based sequential design are presented in Table 1.[8,24,28,31] Costs (step 1b) were estimated using information from the Big CACTUS grant application available before the trial began. The fixed cost of the trial is estimated at £774 809. The variable cost per participant pair (excluding the intervention cost) is £4706. Size of the population (step 1c) affected over 10 years is 215 378. For simplicity, it is assumed there are no costs associated with adopting CSLT into practice if found to be cost-effective. The health economic model and 1000 bootstraps of the CACTUS pilot data estimated (step 1d) the sampling SD of the INMB for a pair of participants at £10 895 and (step 1f) the previous mean and SD for the expected INMB at μ = £3190 and σ = £4118. This positive expected INMB from the previous evidence suggests the CSLT intervention is expected to be cost-effective but with considerable uncertainty.

## *Case Study Step 2: Compute the Optimal Trial Design Within Each of the Design Structures Under Consideration*

The traditional fixed sample size design (a) was undertaken using the original sample size calculation,[28] resulting in $n_a = 95$ participant pairs.

For (b) the value-based design with a fixed sample size, the EVSI was calculated using the standard normal loss function method,[14] and after examining ENBS (EVSI minus trial costs) for many possible sample sizes, the optimal sample size for participant pairs was $n_b = 435$.

For (c), Figure 1 illustrates stopping boundaries for the value-based sequential design with a maximum of $n_c = 95$ (red solid line) and $n_c = 435$ (blue dashed line) participant pairs. The x-axis gives the number of pairs randomized and the y-axis gives the previous/posterior mean of the expected INMB based on accumulating data. The black vertical dotted line shows the point at

**Figure 1.** Calculated stopping boundaries for the value-based sequential design for Big CACTUS trial case study using a maximum sample size of 95 (red solid) and 435 (blue dashed) pairwise allocations.

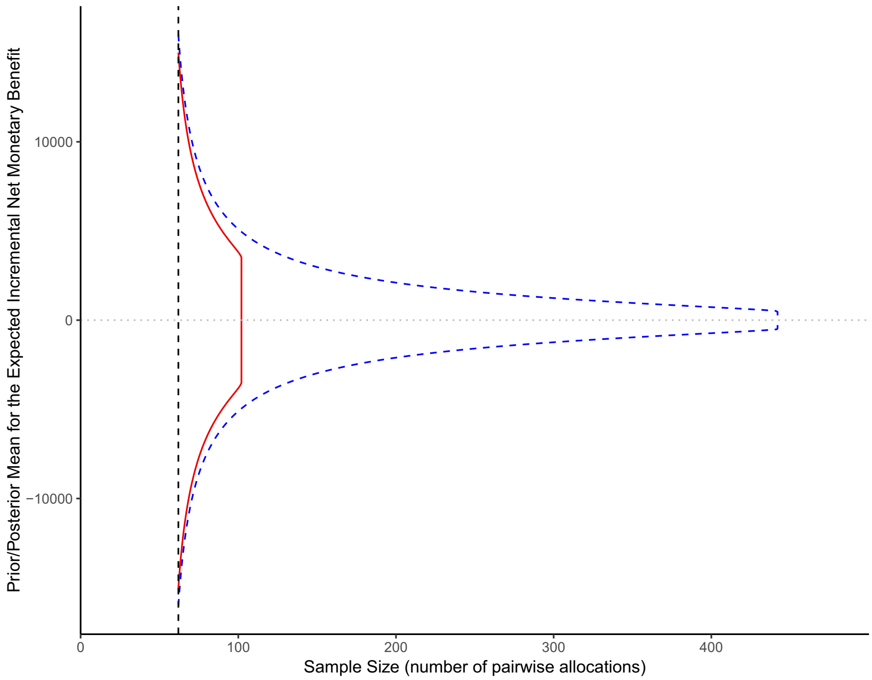
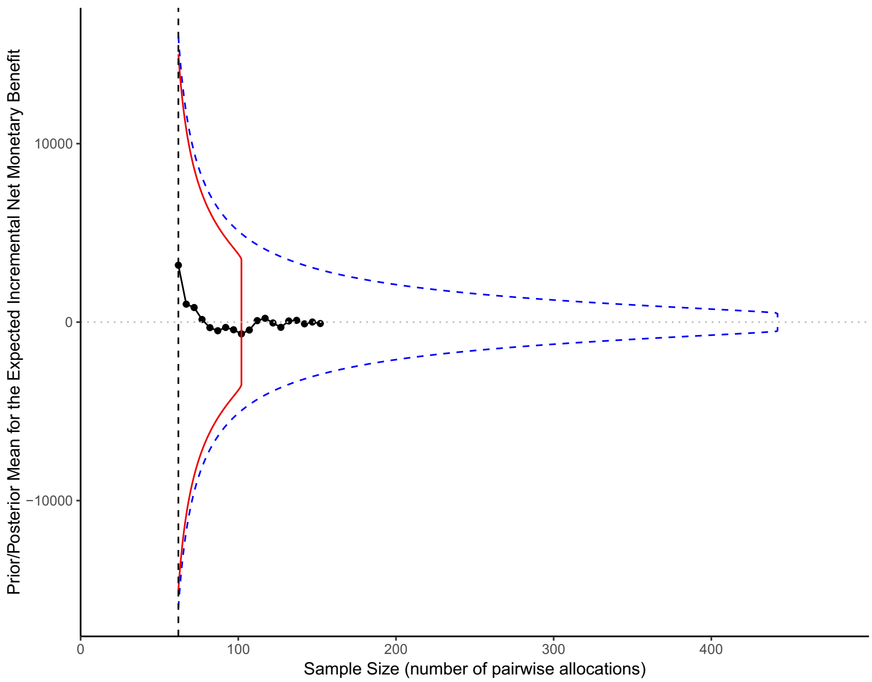



**Table 2.** Summary of characteristics for the 4 designs considered based on information available to the trial team before the trial commences and calculated using htadelay (MATLAB).

| Variable | (a) Original fixed sample size design | (b) Value-based design with optimal fixed sample size | ($c_{n=95}$) Value-based sequential design ($n_c = 95$) | ($c_{n=435}$) Value-based sequential design ($n_c = 435$) |
|---|---|---|---|---|
| Maximum sample size | 95 | 435 | 95 | 435 |
| Maximum total variable cost of the trial | £447 070 | £2 047 110 | £447 070 | £2 047 110 |
| Expected sample size (% increase over original trial) | 95 | 435 | 86.47 | 166.01 |
| Expected total variable cost associated with conducting the proposed trial design | £447 070 | £2 047 110 (+357.9%) | £406 940 (−9.0%) | £781 227 (+74.7%) |
| Population EVSI (% increase over original trial) | £102.4 million | £109.3 million (+6.8%) | £102.6 million (+0.20%) | £109.6 million (+7.0%) |
| Expected net benefit (ENBS) (% increase over original trial) | £101.9 million | £107.3 million (+5.3%) | £102.2 million (+0.3%) | £108.8 million (+6.7%) |
| Proportion of simulated trials that would have stopped early | 0% | 0% | 52.00% | 94.00% |
| Proportion of simulated trials that would have crossed the upper boundary | 0% | 0% | 47.50% | 75.50% |
| Proportion of simulated trials that would have crossed the lower boundary | 0% | 0% | 4.50% | 18.50% |
| Proportion of simulated trials that would not have stopped early | 100% | 100% | 48.00% | 6.00% |
| Proportion of trial results that find CSLT more cost-effective than UC (positive posterior mean incremental net benefit) | | | 77.83% | 77.77% |
| Average over simulations of posterior mean incremental net benefit per person | | | 3064.07 | 3102.75 |

CSLT indicates computerized speech and language therapy; ENBS, expected net benefit of sampling; EVSI, expected value of sample information; UC, usual care.

**Figure 2.** The actual observed Big CACTUS data (black line sample path of posterior INMB calculated) super-imposed upon the stopping boundaries for the value-based sequential design with a maximum sample size of 95 (red solid) and 435 (blue dashed) pairwise allocation. Sample path continues beyond the vertical portion of the solid red line because delays in observing outcomes result in information being updated after the last participant is randomized.



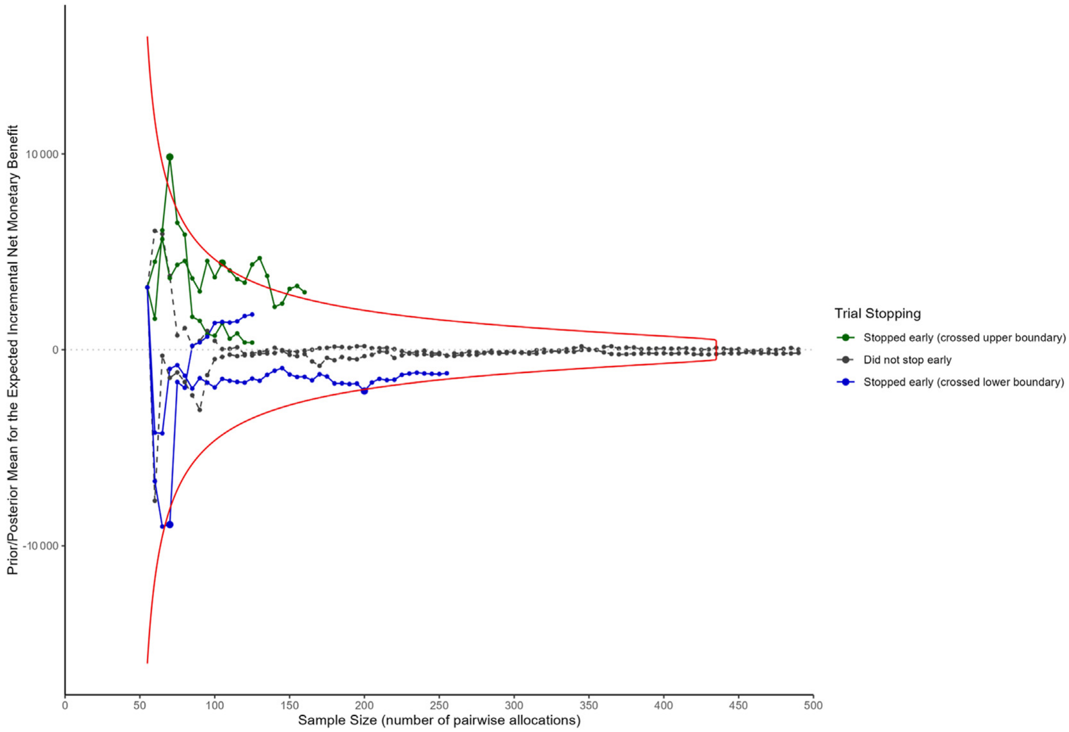

INMB indicates incremental net monetary benefit.

**Table 3.** Summary of the average trial results (5000 simulated trials) in the CACTUS case study comparing the original frequentist fixed sample size design, the value-based fixed sample size design, and the value-based sequential 2-arm design with adaptive stopping with maximum sample size equal to the fixed sample size design, value-based 1-stage design.

| Variable | (a) Original fixed sample size design | (b) Value-based design with optimal fixed sample size | ($c_{n=95}$) Value-based sequential design ($n_c = 95$) | ($c_{n=435}$) Value-based sequential design ($n_c = 435$) |
|---|---|---|---|---|
| Proportion of simulated trials that would have stopped early | 0% | 0% | 9.76% | 31.50% |
| Proportion of simulated trials that would have crossed the upper boundary | 0% | 0% | 7.02% | 5.96% |
| Proportion of simulated trials that would have crossed the lower boundary | 0% | 0% | 2.74% | 25.54% |
| Proportion of simulated trials that would not have stopped early | 100% | 100% | 90.24% | 68.50% |
| Expected sample size (% increase over original trial) | 95 | 435 (+358%) | 92.88 (−2.2%) | 399.57 (+320.6%) |
| Expected total variable cost associated with conducting the proposed trial design | £447 070 | £2 047 110 (+358%) | £437 131 (−2.22%) | £1 880 381 (+321%) |
| Expected total cost (fixed and variable) associated with conducting the proposed trial design | £1 221 879 | £2 821 919 (+131%) | £1 211 940 (−0.813%) | £2 655 190 (+117%) |
| Proportion of trial results that find CSLT more cost-effective than UC (positive posterior mean incremental net benefit) | 35.56% | 9.28% | 36.18% | 13.00% |
| Average over 5000 bootstraps of posterior mean incremental net benefit per person | −130.16 | −371.07 | −108.41 | −328.68 |

*Note.* Also shown is the percentage change relative to the fixed sample size design, for each figure of merit.
CSLT indicates computerized speech and language therapy; UC, usual care.

**Figure 3.** Six simulations of the sample pathways from the bootstrap exercises for a value-adaptive design using the Big CACTUS data with maximum sample size n = 435 participant pairs. Blue and green lines (2 of each) are simulated sample paths, which stopped earlier than maximum sample size. Dark gray lines are 2 simulated sample paths, which continued through to the maximum sample size. The larger dots indicate the sample size where the path crossed the boundary.




which data start to become available from randomized participants accounting for the delay between randomization and data collection. If, at an interim analysis, the value of the posterior mean expected INMB is outside the stopping boundary, then the value-adaptive analysis recommends stopping the trial.

With a maximum sample size of $n_c = 95$ allocations, there is only a brief opportunity to stop the trial early while participants are still being recruited, due to the relatively long follow-up time before health economic outcomes are available compared with the trial recruitment period. This window of opportunity is extended when we consider $n_c = 435$ allocations.

## Case Study Step 3: Comparison of Candidate Designs for the Reimagined Big CACTUS Trial

Step 3 calculates the EVSI and ENBS using only information available before the trial begins. We can also calculate other performance metrics, such as the probability a given arm is selected as best, using parametric Monte Carlo simulations as described in Chick et al.[4]

The estimated expected value of each trial design option over the 10-year decision relevance horizon for the national population is presented in Table 2. We have assumed fixed costs are zero given that they are considered to be the same in all designs (see case study step 1). As expected, none of the trials would stop early when using the fixed and value-based designs. For the value-based sequential design with $n_c = 95$ per arm, 52% of the trials stopped early giving a reduced expected sample size and variable cost compared with the fixed counterpart. Likewise, for the value-based sequential design with $n_c = 435$ per arm, 94% of the trials stopped early giving an expected sample size of 166 and a total variable cost of £781 243.

The estimated expected value of (a) traditional fixed sample size design of $n_a = 95$ per arm was $EVSI_a$ = £102 358 847, with a total variable trial cost $TV_a$ = £447 070 and hence an expected $ENBS_a$ = £101 911 777.

For (b) the fixed sample size value-based design, the resulting optimal sample size was $n_b = 435$ per arm. The $EVSI_b$ = £109 333 687, with an expected variable trial cost $TV_b$ = £2 047 110 and expected $ENBS_b$ = £107 286 577 [+1.8% higher than (a)].

For the value-based sequential designs (c), the expected ENBS result for maximum sample size $n_c = 95$ is $ENBS_{95}$ = £102 208 364 [+0.3% higher than fixed sample size design (a)]. Finally, $ENBS_{435}$ = £ 108 772 817 [+6.7% higher than (a)]. This shows that a value-based sequential design with maximum sample $n_c = 435$ would be expected to produce approximately £7 million more value than the traditional fixed design.

## Case Study Step 4: Using the Observed Big CACTUS Data to Imagine the Trial Conducted With a Value-Based Sequential Design

For step 4 we did not conduct real-time interim analyses of the accruing data because the Big CACTUS trial had already been completed with a traditional fixed design. However, we could imagine the actual trial data accruing. We used these data, together with the health economic model to compute posterior expected INMB of CSLT versus UC after each set of 5 participant pairs. We plot this accruing data "sample path" alongside the stopping boundaries to see whether the trial would have stopped early based on the observed data.

Figure 2 retrospectively compares the actual observed Big CACTUS data with the value-based sequential design stopping boundaries. The Big CACTUS trial data for CSLT and UC were analyzed to compute the (black line) estimated posterior mean for the expected INMB (each dot is after every 5 randomized pairs with reported outcomes). The black line extends beyond the red stopping boundary due to the delay between randomization and outcome observation. The sample path does not cross any stopping boundaries during the trial, only touching the red stopping boundary at its maximum sample size of 95. This indicates that if the original trial had used a value-based sequential design and accrued data as it did, it would not have stopped early. This is due to the small differences in observed QALY and the relatively small cost difference between the interventions. The previous expected INMB was £3190, and as the observed data accrued, the posterior expected INMB estimate fell slightly, and after 20 participant pairs, the estimate was approximately zero and remained within the stopping boundary. At each point of interim analysis, it continued to be cost-effective to collect more data to reduce uncertainty regarding the cost-effectiveness of CSLT versus UC.

### Post-Step 4 Bootstrap Simulation Analysis of Potential Performance of the Value-Adaptive Design

We conducted a simulation analysis to gain insights into each design's performance. The observed Big CACTUS data represents just one realization of what could have happened in the trial. To simulate the possible range of trial results, we took the observed trial data and performed 5000 bootstraps on the actual data, assuming random pair allocation to CSLT or UC. Parameters for the health economic model were calculated at every 5 pairs randomized. Each simulation compared results with stopping boundaries to determine whether the trial would have stopped early. This allowed calculation of expected sample size and total cost for the Big CACTUS trial under a value-based sequential design. The results of these computations are presented in Table 3 and Figure 3. We do not expect the data in Table 2, which are based on data available from the pilot study and which represent an average over uncertainty in the mean INMB, to match the analogous data in Table 3, because the results in Table 3 are computed based on conditioning from the Big CACTUS data that were collected after the pilot study.

The first bootstrap exercise indicates that 9.76% of simulated trials would have stopped early for the value-based sequential design with maximum sample size of $n_c = 95$. With a larger maximum sample size of $n_c = 435$, a second bootstrapping exercise showed 5.66% of trials finishing with a sample size $<95$ (different from the first only due to random noise) and 31.50% finishing before the maximum $n_c = 435$. For the value-based sequential design with $n_c = 95$, the resulting expected sample size is 92.88 (2.2% lower than the maximum 95), indicating that such a trial would rarely stop early. The design with $n_c = 435$ is more likely to stop early, with an expected sample size of 399.57 participant pairs (8.1% lower than the maximum 435, but 320.6% higher than the original fixed design trial n = 95).

This lower expected sample size is reflected in a lower expected total variable. A fixed design trial with a sample size of 435 pairs has a cost of £2.047 million, whereas the value-based sequential design with maximum sample of 435 has an expected total variable cost of £1.880 million.

We note that some of these results differ from the pretrial simulation results calculated using htadelay and summarized in Table 2. This is because the bootstrap results represent a range of possible results from trials that gave the actual data observed, rather than the possible results from the previous distribution (as used in htadelay).

## Discussion

### Summary

Outlined are 4 steps to compare traditional fixed sample size designs, value-based designs with a fixed sample size, and a value-based sequential 2-arm design with adaptive stopping. Using pilot trial data and model-based cost-effectiveness analysis in the case study, it was demonstrated that a value-based sequential design with a maximum sample of $n_c = 435$ is expected to provide better value for money to the funder than shorter or fixed sample size designs. This result is as expected given that the value-based sequential design allows for early termination of the trial if the accumulating evidence indicates one treatment is more cost-effective, enabling a decision on adopting the technology without additional research expenses. Conversely, if data are inconclusive, flexible collection enables gathering more evidence to inform effectiveness and cost-effectiveness decisions with greater precision. The size of the additional value of the value-based sequential design compared with the traditional fixed design in the example was approximately £7 million.

### Contribution to Existing Literature

This case study supplements recent retrospective applications with 2 other UK trials—ProFHER[6] and HERO.[7] It extends the methods to a scenario with pilot data available before the full trial is designed and where a health economic model will be used to analyze the cost-effectiveness adoption decision.

The broader project—Economics of Adaptive Clinical Trials[32]—explored the potential implementation of value-adaptive methods across the portfolio of NIHR-funded research; see reports on stakeholder workshops and case studies.[3,7]

### Implications for Practice and Research

We envision value-adaptive designs becoming part of the repertoire for clinical trials alongside traditional and value-based fixed sample size designs. The framework facilitates design comparisons, suggesting a value-adaptive approach when it significantly enhances expected value, as shown in the example earlier. The framework is adaptable beyond just sample size, including features such as arm allocation probabilities.

Research teams will need to consider the practicality of implementation case by case. Not all trials are suited to adaptive designs,[33] especially when the recruitment period is short but follow-up is long. Practical considerations also include the frequency of interim analyses, the availability of a health economic model before assessing trial designs, and accurate costing of the proposed designs. Multidisciplinary team engagement will be important and has been the focus of qualitative research[21] because stakeholders in health technology assessments will continue to consider clinical effectiveness to be the main focus of a clinical trial.

### Strengths, Limitations, and Further Research

The scale and direction of our case study findings cannot be generalized. The specific costs and benefits of value-based sequential versus fixed designs depend on each case. However, the principles remain. Value-based sequential designs offer the potential to end trials early or continue them longer depending on how the evidence accrues during the trial. Therefore, they can tailor investment of research resources to areas where reduction of uncertainty adds value.

Our methods focus on a publicly funded healthcare system with publicly funded health research investment. Consideration of commercially funded research is beyond the scope of this article.

For simplicity, our case study assumed that the fixed cost and variable costs per participant pair were the same for both fixed sample and the value-based sequential design. In practice, there could be differences because the value-based sequential design could incur additional research costs to cover design and interim analyses.[21]

This application only considered a 2-arm trial. Extending these steps to 3-arm trials is possible by applying multiarm, multistage value-based methods.[22] The methods could also extend to other adaptive designs,[3] including group sequential trials.[34]

## Conclusions

This tutorial outlined 4 steps to calculate the expected value and ENBS for a value-based sequential design in a 2-arm trial with existing pilot data and a cost-effectiveness model. The case study showed that a value-based sequential 2-arm design with an adaptive ability to stop early or collect more data, as appropriate, would be the best design when comparing its expected net benefit with that of the traditional fixed trial design and a value-based fixed design. In addition, further data collection through a longer trial would have represented a cost-effective use of research monies. We hope this tutorial will support increased application of value-based sequential design.

## Author Disclosures

Author disclosure forms can be accessed below in the Supplemental Material section.

## Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2024.06.004.

## Article and Author Information

**Accepted for Publication:** June 3, 2024

**Published Online:** August 1, 2024

doi: https://doi.org/10.1016/j.jval.2024.06.004

**Author Affiliations:** Sheffield Centre for Health and Related Research, The University of Sheffield, Sheffield, England, UK (Flight, Brennan, Wilson); INSEAD, Technology and Operations Management Area, Fontainebleau, France (Chick).

**Correspondence:** Isabelle Wilson, MSc, Sheffield Centre for Health and Related Research (SCHARR), The University of Sheffield, 30 Regent Street, Sheffield, S1 4DA, England, United Kingdom. Email: isabelle.wilson@sheffield.ac.uk

**Author contributions:** *Concept and design:* Flight, Brennan, Chick
*Acquisition of data:* Flight
*Analysis and interpretation of data:* Flight, Brennan, Wilson, Chick
*Drafting of the manuscript:* Flight, Brennan, Wilson, Chick
*Critical revision of the paper for important intellectual content:* Flight, Brennan, Chick
*Statistical Analysis:* Flight, Wilson
*Obtaining funding:* Flight, Brennan
*Administrative, technical, or logistic support:* Wilson
*Supervision:* Brennan, Chick

**Funding/Support:** This work was part of the Economics of Adaptive Clinical Trials project, which was funded by the National Institute for Health Research (NIHR) Clinical Trials Unit Support Funding scheme. The CACTUS pilot trial was funded by the NIHR under its Research for Patient Benefit Programme (grant reference number: PB-PG-1207-14097). This study was also supported by the Stroke and Telehealth themes of the South Yorkshire Collaborations for Leadership in Applied Health Research and Care. NIHR Collaborations for Leadership in Applied Health Research and Care for South Yorkshire acknowledges funding from the NIHR. The study also received support from North of Tyne PCT. The Big CACTUS project was funded by the NIHR Health Technology Assessment Programme (12/21/01). Additional funding support was provided by the Tavistock Trust for Aphasia. Ethics approval for the secondary use of the CACTUS data was granted by the Sheffield Centre for Health and Related Research ethics committee (reference: 032229).

**Role of the Funders/Sponsors:** The views expressed are those of the authors and not necessarily those of the National Health Service, NIHR or the Department of Health and Social Care Health Technology Assessment Programme, the Tavistock Trust for Aphasia, or the Stroke Association.

**Acknowledgment:** The authors thank Martin Forster for his support and expertise in applying the value-based sequential design to the CACTUS case study and providing detailed feedback on the work throughout the project. The authors also thank Munya Dimairo, Steven Julious, Nick Latimer, and Rebecca Palmer for providing comments on early drafts of the work and sharing their expertise from conducting the CACTUS pilot and Big CACTUS trials. Stephen Chick gratefully acknowledges the support of Dr Simba Gill and Sabi Dau to the INSEAD Healthcare Management Initiative.

## REFERENCES

1. Fenwick E, Steuten L, Knies S, et al. Value of information analysis for research decisions—an introduction: report 1 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force. *Value Health*. 2020;23(2):139–150.
2. 12 actions to support and apply research in the NHS. NHS England. https://www.england.nhs.uk/publication/12-actions-to-support-and-apply-research-in-the-nhs/. Accessed April 22, 2024.
3. Flight L, Brennan A, Chick SE, et al. Value-adaptive clinical trial designs for efficient delivery of research–actions, opportunities and challenges for publicly funded trials. White Rose. https://eprints.whiterose.ac.uk/189321/. Accessed April 22, 2024.
4. Chick S, Forster M, Pertile P. A Bayesian decision-theoretic model of sequential experimentation with delayed response. *J R Stat Soc B*. 2017:1439–1462.
5. Pertile P, Forster M, La Torre D. Optimal Bayesian sequential sampling rules for the economic evaluation of health technologies. *J R Stat Soc A*. 2014;177(2):419–438.
6. Forster M, Brealey S, Chick S, et al. Cost-effective clinical trial design: application of a Bayesian sequential model to the ProFHER pragmatic trial. *Clin Trials*. 2021;18(6):647–656.
7. Forster M, Flight L, Corbacho B, et al. Report for the EcoNomics of Adaptive Clinical Trials (ENACT) Project: application of a bayesian value-based sequential model of a clinical trial to the CACTUS and HERO case studies (with guidance material for clinical trials units). White Rose. https://eprints.whiterose.ac.uk/180084/7/enactDeliverable3.pdf. Accessed April 22, 2024.
8. Guide to the methods of technology appraisal 2013. National Institute for Health and Care Excellence. https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781. Accessed April 22, 2024.
9. Flight L, Julious S, Brennan A, Todd S. Expected value of sample information to guide the design of group sequential clinical trials. *Med Decis Making*. 2022;42(4):461–473.
10. Rothery C, Strong M, Koffijberg HE, et al. Value of information analytical methods: report 2 of the ISPOR Value of information analysis emerging good practices task force. *Value Health*. 2020;23(3):277–286.
11. Chick SE, Forster M, Pertile P. Optimal sequential sampling with delayed observations and unknown variance. In: Yilmaz L, Chan WKV, Moon I, Roeder TMK, Macal C, Rossetti MD, eds. *Proceedings of the Winter Simulation Conference*. Piscataway, NJ: IEEE; 2015:3789–3800.
12. Pocock SJ. *Clinical Trials: A Practical Approach*. Chichester, United Kingdom: John Wiley & Sons; 2013.
13. Julious SA. *Sample Sizes for Clinical Trials*. New York, NY: Chapman & Hall/CRC; 2009.
14. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Econ*. 1996;5(6):513–524.
15. Raiffa H, Schlaifer R. *Applied Statistical Decision Theory*. Chichester, United Kingdom: Wiley; 1961.
16. Kunst N, Wilson ECF, Glynn D, et al. Computing the expected value of sample information efficiently: practical guidance and recommendations for four model-based methods. *Value Health*. 2020;23(6):734–742.

17. Heath A, Kunst N, Jackson C, et al. Calculating the expected value of sample information in practice: considerations from 3 case studies. *Med Decis Making*. 2020;40(3):314–326.
18. Alban A, Chick SE, Forster M. Value-based clinical trials: selecting recruitment rates and trial lengths in different regulatory contexts. *Manag Sci*. 2023;69(6):3516–3535.
19. Wilson EC. A practical guide to value of information analysis. *Pharmacoeconomics*. 2015;33(2):105–121.
20. Chick S, Forster M, Pertile P. htadelay package for MATLAB. GitHub. https://github.com/sechick/htadelay. Accessed April 22, 2024.
21. Flight L, Julious S, Brennan A, Todd S, Hind D. How can health economics be used in the design and analysis of adaptive clinical trials? A qualitative analysis. *Trials*. 2020;21(1):1–2.
22. Chick SE, Gans N, Yapar Ö. Bayesian sequential learning for clinical trials of multiple correlated medical interventions. *Manag Sci*. 2022;68(7):4919–4938.
23. Wason JM, Dimairo M, Biggs K, et al. Practical guidance for planning resources required to support publicly funded adaptive clinical trials. *BMC Med*. 2022;20(1):254.
24. Palmer R, Enderby P, Cooper C, et al. Computer therapy compared with usual care for people with long-standing aphasia poststroke: a pilot randomized controlled trial. *Stroke*. 2012;43(7):1904–1911.
25. Palmer R, Dimairo M, Cooper C, et al. Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big Cactus): a multicentre, single-blinded, randomised controlled trial. *Lancet Neurol*. 2019;18(9):821–833.
26. Clinical and cost effectiveness of Aphasia computer therapy compared with usual stimulation or attention control long term post stroke (CACTUS). National Institute for Health and Care Research. https://fundingawards.nihr.ac.uk/award/12/21/01. Accessed April 22, 2024.
27. Microsoft Corporation. Microsoft Excel. https://office.microsoft.com/excel; 2018. Accessed July 30, 2024.
28. Palmer R, Dimairo M, Latimer N, et al. Computerised speech and language therapy or attention control added to usual care for people with long-term post-stroke aphasia: the Big CACTUS three-arm RCT. *Health Technol Assess (Winchester, England)*. 2020;24(19):1.
29. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2023. https://www.R-project.org/. Accessed July 30, 2024.
30. Latimer NR, Bhadhuri A, Alshreef A, et al. Self-managed, computerised word finding therapy as an add-on to usual care for chronic aphasia post-stroke: an economic evaluation. *Clin Rehabil*. 2021;35(5):703–717.
31. Latimer NR, Dixon S, Palmer R. Cost-utility of self-managed computer therapy for people with aphasia. *Int J Technol Assess Health Care*. 2013;29(4):402–409.
32. Value-based Adaptive Clinical Trial Designs for Efficient Delivery of NIHR Research: *EcoNomics of Adaptive Clinical Trials (ENACT)*. University of Sheffield. https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.nihr.ac.uk%2Fdocuments%2Fexplore-nihr%2FEfficient%2520studies%2FSheffield%2520CTRU%2520ENACT%2520-%2520NIHR%2520Final%2520Report%252031March2021.docx&wdOrigin=BROWSELINK. Accessed April 22, 2024.
33. Wason J, Brocklehurst P, Yap C. When to keep it simple–adaptive designs are not always useful. *BMC Med*. 2019;17:1–7.
34. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika*. 1977;64(2):191–199.

## FURTHER READING REFERENCES

1. Stroke_Association. What is aphasia 2021. https://www.stroke.org.uk/what-is-aphasia. Accessed April 22, 2024.
2. RCSLT. RCSLT Resource Manual For Commissioning And Planning Services For SLCN Aphasia - RCSLT 2009 (literature synthesis updated 2014) 2014. https://www.rcslt.org/wp-content/uploads/media/Project/RCSLT/aphasia-resource-updated-feb-2014.pdf. Accessed April 22, 2024.
3. Palmer R, Cooper C, Enderby P, et al. Clinical and cost effectiveness of computer treatment for aphasia post stroke (Big CACTUS): study protocol for a randomised controlled trial. *Trials*. 2015;16:1–12.
4. Spiegelhalter DJ, Freedman LS, Parmar MK. Bayesian approaches to randomized trials. *J R Stat Soc A*. 1994;157(3):357–387.