Moxonidine Trial

Introduction

This report outlines a clinical trial which can be found here: https://pubmed.ncbi.nlm.nih.gov/17968979/.
The trial compared the drug moxonidine against a control in a double-blind trial. We will discuss the trial;
how the sample size was calculated, how the trial was conducted and the results of the trial. We will also
retrospectively compute assurance for this trial, both before and during (in the interim analysis). Using
assurance would have helped the study team in this trial as they would have been able to give distributions
for their beliefs about the control and also for the effect of the experimental treatment. If they were unsure
of their beliefs about either then this could be modelled by the distributions and the assurance calculations
would reflect this. Therefore everyone involved in the trial would be more informed, from the sponsors to
the study team.

Background

Myocardial ischaemia is the leading cause of morbidity and mortality after surgery on patients with coronary
artery disease. Patients were randomly assigned the drug moxonidine or a control (with ratio 1:1) to be
taken the morning before surgery and on days 1 to 4 after surgery. Study medication was administered by a
member of the team blinded to treatment allocation. The endpoint of the trial was measuring the difference
in levels of cTnI before surgery and 7 days post surgery, an increase of 2μg/l of cTnI was interpreted as being
clinically significant in increasing the probability of myocardial ischaemia. Therefore, an increase of 2μg/l
of cTnI is regarded as a failure (not desirable), any increase less than 2μg/l of cTnI is a success (desirable
outcome). An incidence level of 45% was found in a previous investigation for the control and the clinicians
believed the treatment (moxonidine) would decrease incidence levels to 30%.

Trial Design

Based on the above information, a desired Type I error rate α of 0.05 and trial being performed at 80%
power, the necessary sample size was calculated to be 180 patients per group. We can use equation 9.18 in
Julious (2009) to show why the trial chose 180 patients in each arm.

$$n_A = \frac{4}{(\pi_A - \pi_B)^2},$$

Here, we have $\pi_A = 0.45$ and $\pi_B = 0.30$. If we substitute these values into the equation we obtain:

$$
\begin{aligned}
n_A &= \frac{4}{(\pi_A - \pi_B)^2} \\
&= \frac{4}{(0.45 - 0.3)^2} \\
&= \frac{4}{0.0225} \\
&\approx 177.8,
\end{aligned}
$$

so we can see why the value of 180 patients in each arm was chosen.

A blinded interim analysis was prospectively planned after one-third of the calculated study population had been enrolled. Futility was assessed by conditional power under the current trend of the data; if this was calculated to be less than 50% then the study would be considered futile.

Results

At the interim analysis, 141 patients had been recruited into the trial. There were 78 patients in the treatment group and 63 patients in the control group, the results are given below:

	Control (n=63)	Treatment (n=78)
Success	40(63%)	47(60%)
Failure	23(37%)	31(40%)

At the interim analysis, We see that $\approx 40\%$ patients in the treatment group have higher levels of cTnI (higher levels indicate higher levels of myocardial ischaemia), compared to $\approx 37\%$ of patients in the control group. We see that the levels of increased cTnI is greater in the treatment group than the control group, which is clearly not desirable.

This can be analysed with a χ^2 test, as seen below.

```
MoxonidineData <- data.frame(RaisedcTni = c(23, 31), NotRaisedcTni = c(40, 47),
                             row.names = c("Control", "Treatment"))
chisq.test(MoxonidineData, correct = F)
```

```
##
##  Pearson's Chi-squared test
##
## data:  MoxonidineData
## X-squared = 0.15441, df = 1, p-value = 0.6944
```

We can see that the p-value calculated is 0.694, which clearly is not significant at all. We can calculate conditional power here.

Conditional power

We will assume different scenarios here. The first scenario is assuming that the remaining data will follow the same pattern of the data accrued so far. We see that when we assume this, the conditional power drops to 0.

```
## [1] 0
```

In the following scenario, we assume the remaining data will follow that of the clinicians beliefs before the trial started: control will have an incidence rate of 45% and treatment will have a lower incidence rate of 30%. Even in this very optimistic scenario (as the data so far does not support these beliefs), the conditional power is very low. This suggests that the trial should be stopped for futility.

```
## [1] 0.027
```

Assurance for this trial

Trial design stage

The design of this trial is very traditional, very frequentist. We have a placebo (for which we have a point-estimate), a treatment (for which we have an estimate of its effects in comparison to the placebo), we have the type I error rate ($\alpha = 0.05$) and we have some power that we wish to perform the trial at (here, $1 - \beta = 0.8$). As a result, we saw how when we substitute these values into the equations, we obtain 180 patients in both of the two arms.

However, this method is sub-optimal for a number of reasons. The first reason is that this calculation assumes that the treatment is in fact better than the placebo - which we cannot possibly know as we would not be doing the trial otherwise. Therefore, the power of the test (80%) is the conditional probability that we will find an effect, if such an effect exists. In reality, there will be times where the treatment is no better than the placebo (as was the case in this trial) so the power is a misrepresentation of how effective a trial actually is.

The second reason is that we only have point-estimates for the placebo and the treatment effect. From historical data, we should be able to construct a distribution for the effectiveness of the placebo, and from possible phase II trials, pilot studies and in collaboration with clinical investigators, we should be able to construct a distribution for the effectiveness of the treatment. Even if we are not very confident about our estimates in the effectiveness of the treatment, these prior beliefs can be incorporated into the distribution.

Trial setup

The outcome is a binary variable, we have r_i successes (note: a 'success' here is **not** a desirable outcome) out of n_i patients for both treatments ($i = 1$ for the placebo, $i = 2$ for the treatment). Therefore, the hypothesis test we conduct at the end of the trial is:

$$H_0 = \theta_1 = \theta_2$$
$$H_1 = \theta_1 \neq \theta_2$$

Therefore, to calculate assurance methods for this trial design, we require prior beliefs regarding θ_1 and θ_2. From the information in the trial, we know that $E[\theta_1] = 0.45$. The trial investigators will be able to calculate the SD of θ_1 accurately, but for illustration, we will choose $\text{sd}[\theta_1] = 0.1$. We can find a Beta distribution (we use a Beta distribution as it only takes values [0,1] which is desirable for a proportion) which captures these beliefs by:

```
estBetaParams <- function(mu, var) {
  alpha <- ((1 - mu) / var - 1 / mu) * mu ^ 2
  beta <- alpha * (1 / mu - 1)
  return(params = list(alpha = alpha, beta = beta))
}

estBetaParams(0.45, 0.1^2)
```

```
## $alpha
## [1] 10.6875
##
## $beta
## [1] 13.0625
```

So we can see that these beliefs can be approximately captured by a Be(10.7, 13.1) distribution.

To elicit beliefs about θ_2, we can instead think about the difference between the control treatment and the experimental treatment. Let $\rho = \theta_1 - \theta_2$. We prefer to think about the difference as θ_1 and θ_2 are not independent. From the trial setup, we know that the clinicians believe that the experimental treatment has an effect size of 33%, this indicates that $E[\theta_2] = 0.3$ and $E[\rho] = 0.15$. Therefore, we can model $\rho \sim N(0.15, \nu)$. Again, the trial investigators will be able to choose an appropriate value for ν, so for illustration we will calculate the assurance under three different scenarios (note: the third scenario has different value for the mean):
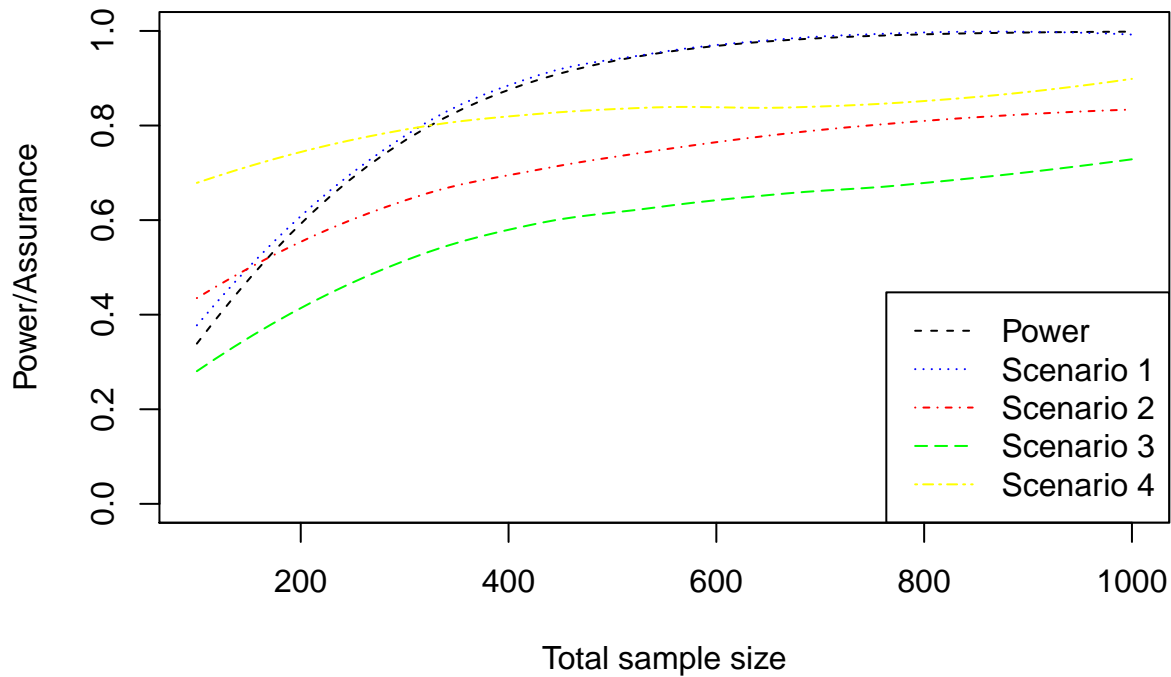
Scenario 1: $\rho \sim N(0.15, 0.0001)$

Scenario 2: $\rho \sim N(0.15, 0.01)$

Scenario 3: $\rho \sim N(0.10, 0.01)$

Scenario 4: $\rho \sim Unif(\theta_1 - 1, \theta 1)$

We can calculate assurance under the four scenarios. This can be seen in the figure below - alongside the traditional power curve.

In scenario 1, the prior distribution for ρ indicates a strong belief that the treatment is 33% more effective than the control. This is shown by the assurance essentially being the power curve (where we assume the treatment is better).

In scenario 2, the prior for ρ indicates that the beliefs are not quite as strong as in scenario 1. This is modelled through the value of ν being larger and shown in the plot by the assurance being lower (reaches a maximum of around 80%)

In scenario 3, the prior for ρ indicates that the treatment may not be as effective as 33%. This is shown in the plot as the assurance is lower than both scenarios 1 and 2.

Interim analysis stage

We can also perform assurance calculations during the trial. This feels natural in a Bayesian paradigm; we have prior distributions (about θ_1 and ρ) and some data, which we can combine together to obtain a posterior distribution.

The following four plots show the four different scenarios; in each plot we have the power, the assurance at the design stage and the assurance at the interim analysis stage. The total sample sizes only reach a maximum of 360 (180 in each arm) here, as this was the number of patients planned to be enrolled in the trial.

```
data {
  int<lower=0> cE;                // control events
  int<lower=0> cN;                // number of patients in control group
  int<lower=0> tE;                // treatment events
  int<lower=0> tN;                // number of patients in treatment group
}

parameters {
  real<lower=0, upper=1> theta1; // chance of success in control group
  real rho;                       // difference between treatment and control (theta1-theta2)
}

transformed parameters{
  real<lower=0, upper=1> theta2; // chance of success in treatment group
  theta2 = theta1 - rho;
}

model {
  theta1 ~ beta(10.7, 13.1);
  rho ~ normal(0.15, 0.01);
  cE ~ binomial(cN, theta1);
  tE ~ binomial(tN, theta2);
}

generated quantities {
  int<lower=0> controlevents;
  int<lower=0> treatmentevents;
  controlevents = binomial_rng(117, theta1);
  treatmentevents = binomial_rng(102, theta2);
}
```

```
##
## SAMPLING FOR MODEL '5ad629cb87cba5efe789b1a5dff7c1bc' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 2e-05 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.2 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 10000 [  0%]  (Warmup)
## Chain 1: Iteration: 1000 / 10000 [ 10%]  (Warmup)
## Chain 1: Iteration: 2000 / 10000 [ 20%]  (Warmup)
## Chain 1: Iteration: 3000 / 10000 [ 30%]  (Warmup)
```

```
## Chain 1: Iteration: 4000 / 10000 [ 40%]  (Warmup)
## Chain 1: Iteration: 5000 / 10000 [ 50%]  (Warmup)
## Chain 1: Iteration: 5001 / 10000 [ 50%]  (Sampling)
## Chain 1: Iteration: 6000 / 10000 [ 60%]  (Sampling)
## Chain 1: Iteration: 7000 / 10000 [ 70%]  (Sampling)
## Chain 1: Iteration: 8000 / 10000 [ 80%]  (Sampling)
## Chain 1: Iteration: 9000 / 10000 [ 90%]  (Sampling)
## Chain 1: Iteration: 10000 / 10000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 0.06953 seconds (Warm-up)
## Chain 1:                0.066047 seconds (Sampling)
## Chain 1:                0.135577 seconds (Total)
## Chain 1:
```