MFIT Sample Size

Trial Operating Characteristics

James Totterdell

2022-04-11

Contents

1 Fixed-Sample Size Power

Ignoring repeated measures, a sample size of 100 participants per arm, assuming drop-out of 20%, would provide power 0.8 for an effect size of $\Delta = \mu_1 - \mu_0 = 0.45\sigma$. For example, assuming that FACIT-Fatigue scores are distributed according to $Y_j \sim (\mu_j, 5^2)$, then the sample size would have power 0.8 for an increase of 2.25 in group means from control to active treatment.

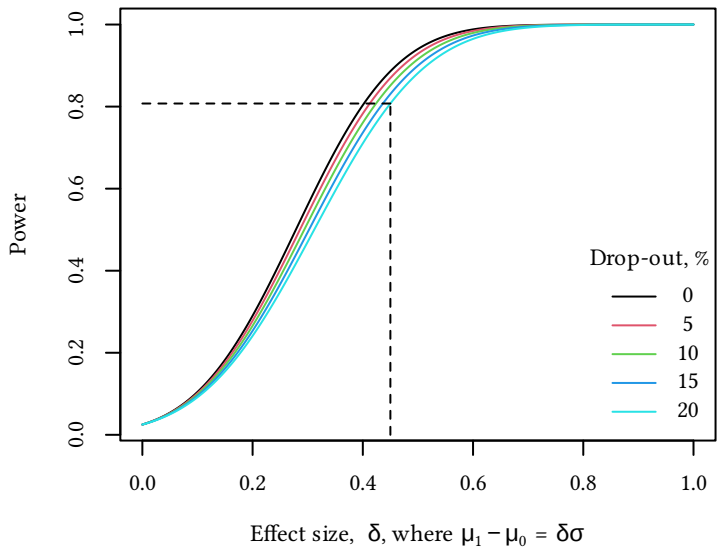

Figure 1: Fixed sample size power for two sample t-test of size $\alpha = 0.025$ with $n = 100$ per arm, assuming varying drop-out and effect size, δ, relative to the standard deviation σ.

2 Group-Sequential Design

The study consists of four arms: one control group and three active interventions. Accrual is assumed to be on average 3 participants per week up to a maximum of 400 participants.

Denote the treatment group means at each time-point by $\mu_{jt},\ j = 0, 1, 2, 3$ for $t = 0, 1, 2, 3$ corresponding to baseline, 4, 8, and 12 weeks after randomisation. The primary endpoint is at 12-weeks after randomisation.

The primary quantity of interest is the difference in FACIT-Fatigue at 12-weeks for each active treatment group relative to the control group, that is,

$$\Delta_j = \mu_{j,3} - \mu_{0,3}, \quad j = 1, 2, 3.$$

Another quantity of interest is the relative effectiveness of each active treatment compared to each other active treatment.

$$\delta_j = \mu_{j,3} - \max_{j' \neq j} \mu_{j',t}, \quad j = 1, 2, 3.$$

We define an active treatment to be best if mean FACIT-Fatigue is maximal (amongst all active treatments) under that treatment, that is, $\delta_j > 0$, and we define an active treatment to be effective, if it has higher mean FACIT-Fatigue compared to control, $\Delta_j > 0$.

The truth of these assertions are unknown, but we quantify their probability under the assumed model according to

$$\omega_j = \Pr(\Delta_j > 0 | \text{data})$$
$$\pi_j = \Pr(\delta_j > 0 | \text{data})$$

such that ω_j is the posterior probability that treatment j is effective, and π_j is the posterior probability that treatment j is best amongst the active treatments.

2.1 Model

The analysis model used in the simulations ignored the longitudinal outcomes, using was

$$\begin{aligned}
y_{i,3} | \alpha, \beta, \sigma &\sim \text{Normal}(\alpha + x^{\mathsf{T}}_{\texttt{trt}[i]}\beta, \sigma^2) \\
\mu_{j,3} &= \alpha + x^{\mathsf{T}}_j \beta \\
\alpha &\sim \text{Normal}(40, 5) \\
\beta &\sim \text{Normal}(0, 5) \\
\sigma &\sim \text{Half-}t(3, 0, 5)
\end{aligned}$$

where $x_{\texttt{trt}[i]}$ for $\texttt{trt}[i] \in \{0, 1, 2, 3\}$ denotes the treatment design vector corresponding to participants i's assigned treatment.

2.2 Interim Analyses

Interim analyses are scheduled to occur when 100, 200, 300 and 400 participants reach their primary endpoint.

At each interim analysis, we pre-specify a number of decision rules in terms of the previously defined

quantities. These decision rules may be to drop specific treatment arms, or to stop the trial altogether. We denote by J' the number of available arms at the time of the interim analysis, including control. At the first interim analysis, $J' = 4$.

- **Effective**: if $\omega_j > \epsilon_0$, then treatment j has probability greater than ϵ_0 of being effective compared to control.
- **Ineffective**: if $\omega_j < 1 - \epsilon_1$ then treatment j has probability greater than ϵ_1 of being ineffective compared to control, and therefore may be dropped from the trial.
- **Superiority**: if $\pi_j > \epsilon_2$, then treatment j has probability greater than ϵ_1 of being the best active treatment, and therefore, all other active treatments may be dropped from the trial.
- **Inferiority**: if $\pi_j < \epsilon_2/(J' - 2)$, then treatment j has probability less than $\epsilon_1/(J' - 2)$ of being the best active treatment, and therefore may be dropped from the trial.
- **Stopping**: the trial may be stopped early if one active treatment is superior and effective.

Note that, once a treatment has been dropped, we do not allow it to come back into the trial. Therefore, once a treatment has been declared harmful, inferior, or superior, then that assertion is assumed to hold true for the remainder of the trial. An arm which has been dropped is not considered to possibly be the best arm, and therefore, is excluded from future calculations of π_j. That is, we enforce $\pi_j = 0$ for that arm.

Effectiveness may change throughout of the course of the trial, for example, an intervention may meet the threshold following the first analysis, but no longer does following the second analysis, etc.

2.3 Response-Adaptive Randomisation

Initially, each treatment arm are allocated with equal-probability of $r_j = 1/J'$ where $J' = 4$. The probability of allocation to the control group is fixed throughout the trial to be equal to the reciprical of the number of arms still on study, $r_0 = 1/J'$.

For the remaining arms, the allocation may be zero if the arm has been dropped, or it may be proportional to the probability it is the best active treatment.

Following each interim analysis, the allocation probability of the 3 active arms are updated according to

$$r_j = \begin{cases} 0 & \text{if the arm has been dropped} \\ \frac{(\pi_j/n_j)^k}{\sum_{j=1}^{3}(\pi_j/n_j)^k} & \text{otherwise} \end{cases}, \quad j = 1, 2, 3.$$

for some scaling factor k.

Note that if $k = 0$, then $r_j = 1/J'$ for each arm and if $k = 1$ then $r_j = \pi_j$ (assuming equal sample sizes allocated).

This achieves two arms: more participants are expected to receive the best active intervention, and increased information for the best active intervention.

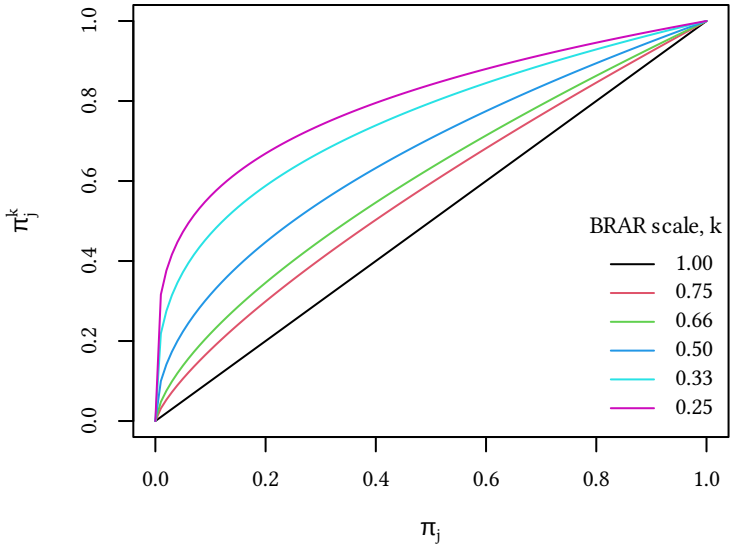

Figure 2: Scaling of probability best for use in RAR. Note small values are upweighted more as k decreases towards 0.

3 Group-sequential Operating Characteristics

The simulations presented here assume:

- expected accrual of 3 participants per week according to homogeneous Poisson process, so at each interim expect ≈ 36 enrolled but without follow-up.
- interim analysis when 100, 200, 300, and 400 participants reach their primary endpoint.
- 20% drop-out, so that, for example when 100 participants reach their primary endpoint, 20 have missing 12-week FACIT-Fatigue data.
- decision thresholds of $\epsilon_0 = \epsilon_1 = \epsilon_2 \in \{0.975, 0.98, 0.99\}$.
- ignores repeated measures analysing 12-week FACIT-Fatigue as response in a linear model conditional on treatment group

Table 1: Summary of trial operating characteristics, n = 10,000 simulations, $\epsilon_0 = \epsilon_1 = \epsilon_2 = 0.98$.

Effect	Number	Treatment	Allocated	Superior	Inferior	Effective	Ineffective	Active
0	0	0 - control	83					
		1	79	0.01	0.10	0.02	0.05	0.90
		2	78	0.01	0.10	0.02	0.05	0.89
		3	79	0.01	0.10	0.02	0.05	0.89
0.5SD	1	0 - control	71					
		1	93	0.83	0.00	0.92	0.00	1.00
		2	46	0.00	0.89	0.01	0.04	0.11
		3	46	0.00	0.89	0.01	0.04	0.11
	2	0 - control	89					
		1	92	0.05	0.05	0.86	0.00	0.95
		2	93	0.05	0.05	0.87	0.00	0.95
		3	37	0.00	0.94	0.01	0.04	0.06
	3	0 - control	82					
		1	78	0.01	0.08	0.79	0.00	0.92
		2	78	0.01	0.08	0.78	0.00	0.92
		3	79	0.01	0.08	0.80	0.00	0.92
1SD	1	0 - control	37					
		1	42	1.00	0.00	1.00	0.00	1.00
		2	31	0.00	1.00	0.01	0.03	0.00
		3	31	0.00	1.00	0.01	0.03	0.00
	2	0 - control	92					
		1	93	0.06	0.06	0.99	0.00	0.94
		2	92	0.06	0.06	1.00	0.00	0.94
		3	30	0.00	1.00	0.02	0.05	0.00
	3	0 - control	81					
		1	78	0.01	0.09	1.00	0.00	0.91
		2	79	0.01	0.09	1.00	0.00	0.91
		3	79	0.01	0.09	1.00	0.00	0.91

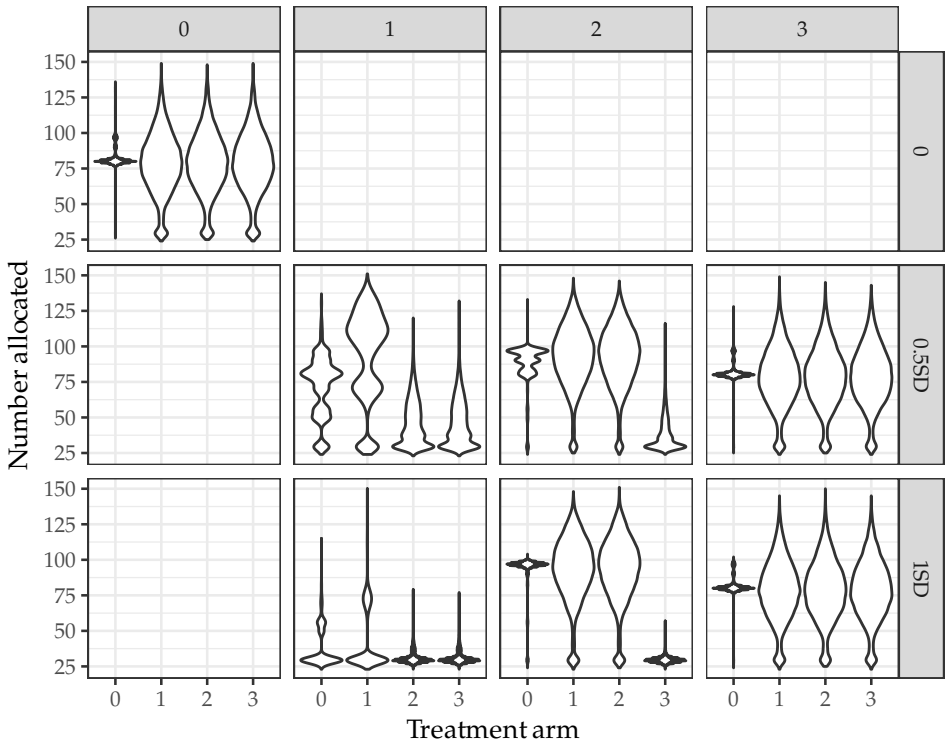

Figure 3: Distribution of number of participants allocated to each treatment by effect size and number of affected treatments.

Table 2: Summary of trial operating characteristics, n = 10,000 simulations, $\epsilon_0 = \epsilon_1 = \epsilon_2 = 0.975$.

Effect	Number	Treatment	Allocated	Superior	Inferior	Effective	Ineffective	Active
0	0	0 - control	84					
		1	78	0.02	0.12	0.02	0.06	0.87
		2	78	0.02	0.12	0.02	0.06	0.87
		3	78	0.02	0.12	0.02	0.06	0.87
0.5SD	1	0 - control	70					
		1	89	0.86	0.00	0.93	0.00	1.00
		2	45	0.00	0.91	0.02	0.05	0.09
		3	45	0.00	0.92	0.02	0.05	0.09
	2	0 - control	88					
		1	92	0.06	0.06	0.87	0.00	0.94
		2	91	0.06	0.06	0.87	0.00	0.94
		3	37	0.00	0.95	0.02	0.06	0.05
	3	0 - control	82					
		1	78	0.01	0.10	0.81	0.00	0.90
		2	78	0.01	0.10	0.81	0.00	0.90
		3	79	0.01	0.09	0.81	0.00	0.91
1SD	1	0 - control	36					
		1	40	1.00	0.00	1.00	0.00	1.00
		2	31	0.00	1.00	0.02	0.05	0.00
		3	31	0.00	1.00	0.01	0.04	0.00
	2	0 - control	91					
		1	91	0.07	0.07	0.99	0.00	0.93
		2	91	0.07	0.07	0.99	0.00	0.93
		3	30	0.00	1.00	0.01	0.06	0.00
	3	0 - control	82					
		1	78	0.01	0.11	1.00	0.00	0.89
		2	78	0.01	0.10	1.00	0.00	0.90
		3	78	0.01	0.10	1.00	0.00	0.90

Table 3: Summary of trial operating characteristics, n = 10,000 simulations, $\epsilon_0 = \epsilon_1 = \epsilon_2 = 0.99$.

Effect	Number	Treatment	Allocated	Superior	Inferior	Effective	Ineffective	Active
0	0	0 - control	82					
		1	79	0.00	0.06	0.01	0.03	0.94
		2	80	0.00	0.05	0.01	0.02	0.94
		3	79	0.00	0.06	0.01	0.02	0.94
0.5SD	1	0 - control	75					
		1	103	0.73	0.00	0.88	0.00	1.00
		2	47	0.00	0.82	0.01	0.02	0.18
		3	47	0.00	0.82	0.01	0.02	0.18
	2	0 - control	88					
		1	95	0.03	0.03	0.81	0.00	0.97
		2	94	0.02	0.03	0.81	0.00	0.97
		3	39	0.00	0.88	0.01	0.03	0.12
	3	0 - control	81					
		1	79	0.00	0.05	0.72	0.00	0.95
		2	79	0.00	0.04	0.72	0.00	0.96
		3	80	0.00	0.05	0.72	0.00	0.95
1SD	1	0 - control	41					
		1	48	1.00	0.00	0.99	0.00	1.00
		2	31	0.00	1.00	0.01	0.02	0.00
		3	31	0.00	1.00	0.01	0.02	0.00
	2	0 - control	94					
		1	95	0.03	0.03	1.00	0.00	0.97
		2	95	0.03	0.03	0.99	0.00	0.97
		3	30	0.00	1.00	0.01	0.03	0.00
	3	0 - control	81					
		1	80	0.00	0.05	0.99	0.00	0.95
		2	79	0.00	0.05	1.00	0.00	0.95
		3	79	0.00	0.05	1.00	0.00	0.95

3.1 Stricter RAR

Table 4: Summary of trial operating characteristics, n = 10,000 simulations, $\epsilon_0 = \epsilon_1 = \epsilon_2 = 0.98$, BRAR scale factor $k = 0.25$.

Effect	Number	Treatment	Allocated	Superior	Inferior	Effective	Ineffective	Active
0	0	0 - control	83					
		1	78	0.01	0.11	0.02	0.05	0.89
		2	78	0.01	0.10	0.02	0.05	0.89
		3	79	0.01	0.10	0.02	0.05	0.89
0.5SD	1	0 - control	71					
		1	83	0.85	0.00	0.91	0.00	1.00
		2	50	0.00	0.90	0.01	0.04	0.10
		3	50	0.00	0.91	0.01	0.05	0.09
	2	0 - control	88					
		1	90	0.05	0.05	0.87	0.00	0.95
		2	91	0.05	0.05	0.87	0.00	0.95
		3	40	0.00	0.96	0.01	0.05	0.04
	3	0 - control	82					
		1	78	0.01	0.09	0.81	0.00	0.91
		2	79	0.01	0.08	0.80	0.00	0.92
		3	79	0.01	0.08	0.81	0.00	0.92
1SD	1	0 - control	37					
		1	40	1.00	0.00	1.00	0.00	1.00
		2	32	0.00	1.00	0.01	0.03	0.00
		3	32	0.00	1.00	0.01	0.03	0.00
	2	0 - control	92					
		1	92	0.06	0.06	0.99	0.00	0.94
		2	93	0.06	0.06	0.99	0.00	0.94
		3	30	0.00	1.00	0.01	0.06	0.00
	3	0 - control	82					
		1	78	0.01	0.09	1.00	0.00	0.91
		2	79	0.01	0.09	1.00	0.00	0.91
		3	78	0.01	0.09	1.00	0.00	0.91

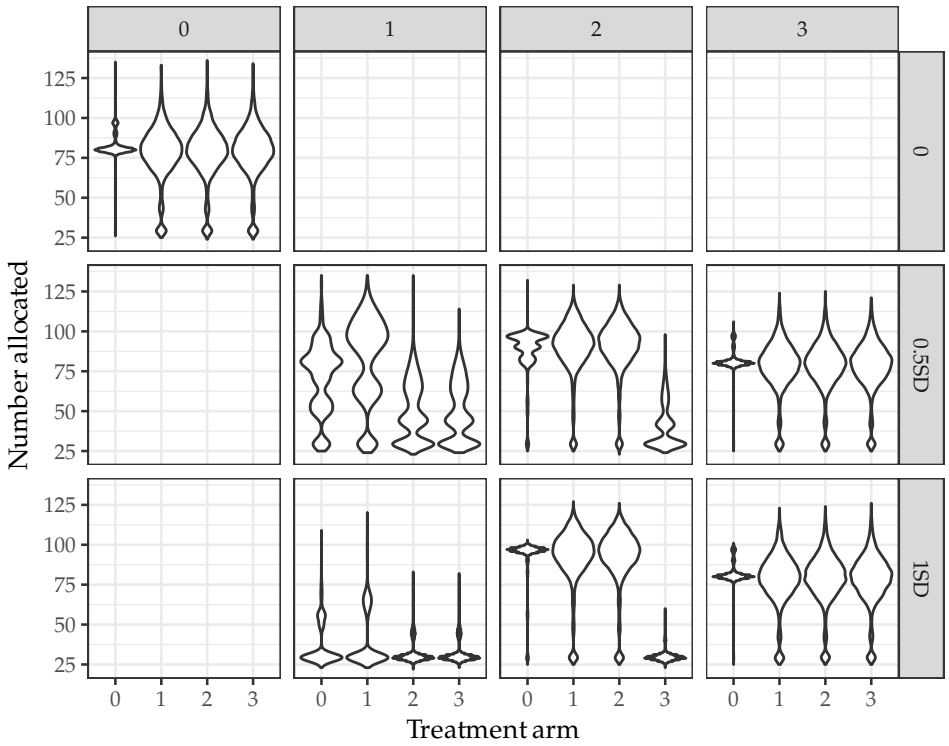

Figure 4: Distribution of number of participants allocated to each treatment by effect size and number of affected treatments.

3.2 Flexible Boundary

Table 5: Summary of trial operating characteristics, n = 10,000 simulations, $\epsilon_0 = \epsilon_1 = \epsilon_2 = 0.9^{\sqrt{t}}$.

Effect	Number	Treatment	Allocated	Superior	Inferior	Effective	Ineffective	Active
0	0	0 - control	83					
		1	78	0.01	0.11	0.02	0.06	0.88
		2	79	0.01	0.10	0.02	0.05	0.89
		3	79	0.01	0.09	0.02	0.04	0.91
0.5SD	1	0 - control	71					
		1	94	0.87	0.00	0.93	0.00	1.00
		2	45	0.00	0.92	0.02	0.04	0.08
		3	45	0.00	0.92	0.01	0.04	0.08
	2	0 - control	88					
		1	93	0.06	0.05	0.85	0.00	0.95
		2	92	0.05	0.06	0.87	0.00	0.94
		3	38	0.00	0.94	0.01	0.04	0.06
	3	0 - control	82					
		1	79	0.02	0.09	0.81	0.00	0.91
		2	79	0.01	0.09	0.81	0.00	0.91
		3	78	0.01	0.09	0.81	0.00	0.91
1SD	1	0 - control	39					
		1	45	1.00	0.00	1.00	0.00	1.00
		2	31	0.00	1.00	0.01	0.03	0.00
		3	31	0.00	1.00	0.02	0.02	0.00
	2	0 - control	93					
		1	92	0.07	0.05	1.00	0.00	0.95
		2	93	0.05	0.07	0.99	0.00	0.93
		3	30	0.00	1.00	0.02	0.06	0.00
	3	0 - control	82					
		1	80	0.01	0.08	1.00	0.00	0.92
		2	79	0.01	0.07	1.00	0.00	0.93
		3	78	0.01	0.10	0.99	0.00	0.90

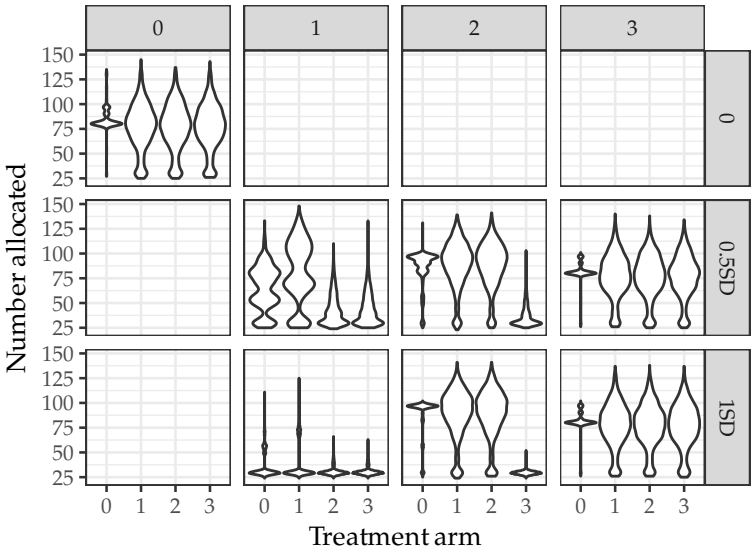

Figure 5: Distribution of number of participants allocated to each treatment by effect size and number of affected treatments.

Table 6: Summary of trial operating characteristics, n = 10,000 simulations, $\epsilon_0 = \epsilon_1 = \epsilon_2 = 0.95^{\sqrt{t}}$.

Effect	Number	Treatment	Allocated	Superior	Inferior	Effective	Ineffective	Active
0	0	0 - control	85					
		1	79	0.03	0.16	0.05	0.08	0.83
		2	77	0.04	0.19	0.05	0.09	0.80
		3	77	0.03	0.19	0.05	0.10	0.79
0.5SD	1	0 - control	65					
		1	80	0.93	0.00	0.94	0.00	1.00
		2	43	0.00	0.96	0.03	0.08	0.04
		3	44	0.00	0.95	0.02	0.07	0.05
	2	0 - control	87					
		1	90	0.10	0.10	0.90	0.00	0.91
		2	89	0.09	0.10	0.91	0.00	0.90
		3	36	0.00	0.98	0.03	0.08	0.02
	3	0 - control	82					
		1	78	0.02	0.14	0.87	0.00	0.86
		2	77	0.02	0.16	0.85	0.00	0.84
		3	77	0.02	0.15	0.84	0.00	0.85
1SD	1	0 - control	35					
		1	38	1.00	0.00	0.99	0.00	1.00
		2	31	0.00	1.00	0.02	0.06	0.00
		3	30	0.00	1.00	0.01	0.06	0.00
	2	0 - control	89					
		1	90	0.12	0.12	1.00	0.00	0.88
		2	90	0.12	0.12	0.99	0.00	0.88
		3	30	0.00	1.00	0.04	0.11	0.00
	3	0 - control	82					
		1	78	0.02	0.13	1.00	0.00	0.87
		2	77	0.02	0.14	1.00	0.00	0.86
		3	79	0.02	0.14	1.00	0.00	0.86

Table 7: Summary of trial operating characteristics, n = 10,000 simulations, $\epsilon_0 = \epsilon_1 = \epsilon_2 = 0.975^{\sqrt{t}}$.

Effect	Number	Treatment	Allocated	Superior	Inferior	Effective	Ineffective	Active
0	0	0 - control	87					
		1	75	0.07	0.29	0.10	0.18	0.68
		2	74	0.07	0.29	0.10	0.19	0.67
		3	75	0.08	0.27	0.08	0.16	0.68
0.5SD	1	0 - control	58					
		1	68	0.98	0.00	0.97	0.00	1.00
		2	41	0.00	0.98	0.06	0.14	0.01
		3	42	0.00	0.99	0.05	0.12	0.01
	2	0 - control	82					
		1	84	0.20	0.17	0.91	0.00	0.83
		2	83	0.17	0.20	0.92	0.00	0.80
		3	35	0.00	0.99	0.07	0.15	0.01
	3	0 - control	81					
		1	75	0.05	0.24	0.90	0.00	0.76
		2	74	0.05	0.23	0.90	0.00	0.77
		3	75	0.06	0.23	0.90	0.00	0.77
1SD	1	0 - control	33					
		1	34	1.00	0.00	1.00	0.00	1.00
		2	30	0.00	1.00	0.03	0.09	0.00
		3	30	0.00	1.00	0.05	0.12	0.00
	2	0 - control	84					
		1	84	0.19	0.19	0.99	0.00	0.81
		2	83	0.19	0.19	1.00	0.00	0.81
		3	30	0.00	1.00	0.07	0.17	0.00
	3	0 - control	81					
		1	75	0.05	0.27	1.00	0.00	0.73
		2	75	0.07	0.24	1.00	0.00	0.76
		3	74	0.06	0.26	1.00	0.00	0.74