CLARITY 2.0 Simulations

James Totterdell and Michael Dymock

2021-11-04

Contents

1 Background

1.1 Aim

The aim of CLARITY 2.0 is to evaluate the safety and efficacy of dual treatment with repagermanium and candesartan compared to placebo for patients hospitalised for management of COVID-19.

1.2 Objectives

The primary objective is to evaluate the safety and efficacy of dual treatment with repagermanium and candesartan in patients hospitalised with COVID-19 disease, assessed by clinical Health Score at day 14.

The clinical health score is a 8 (or still 7?) point ordinal scale:

1. Not hospitalised, no limitations on activities. ...
2. Death.

1.3 Interventions

Participants will be randomised into three treatment arms:

1. Candesartan + repagermanium, (Investigational Arm)
2. Candesartan + placebo [repagermanium], (Control Arm #1)
3. Placebo [candesartan] + placebo [repagermanium] (Control Arm #2)

1.4 Randomisation

Initially, treatment allocation will be with a 1:1:1 block randomisation between the three arms, stratified by centre.

1.5 Sample Size

Stage 1 - up to 600 participants in India only, following which one control group could potentially be dropped Stage 2 - up to 1,500 additional participants between 3 or 2 arms (do we want this total sample size to be calculated on the basis of the stage 1 data?, i.e. use the posterior from stage 1 data to decide planned sample size for stage 2, or just stick with 1,500?).

Total sample size may therefore be up to 2,100 participants.

After stage 1, interim analyses will be conducted every 300 participants with follow-up where effectiveness will be assessed. If results appear futile then the trial may stop early for lack of effectiveness.

1.6 Statistical Analysis

A cumulative logistic regression model will be used to assess the effectiveness of the investigational arm compared to each control arm (or the single control arm). The investigational arm will be compared to each control arm, and effectiveness assessed in terms of the common odds ratio of being in a better clinical health status at day 14. A 95% credible interval and the posterior probability that the common odds ratio is less than zero will be reported.

1.7 Decision Rules

The trial will only stop for effectiveness if the investigational arm is found superior to either the single (two arm), or both continuing control arms (three arm).

2 Approximate Power

As an indication, consider two arms with 1:1 assignment and a sample size of 200 per arm (Stage 1). Assume that there is no missingness and no drop-out.

The power function of an approximate test of equality with $\alpha = 0.05$ is given in Figure 1 assuming that the distribution of the levels of the outcome in the control group is (0.16, 0.29, 0.32, 0.13, 0.02, 0.01, 0.01, 0.06).

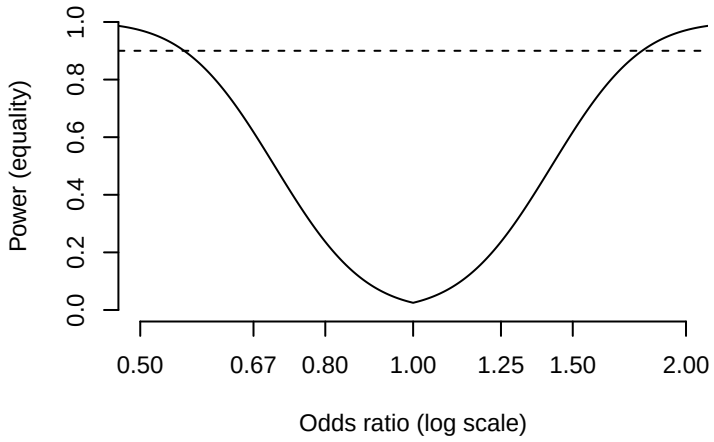

Figure 1: Two arm with 200 per arm.

Figure 2 presents the same, but assuming 700 participants per arm (i.e Phase III continues with three arms up to 2,100 inclusive of Phase II).

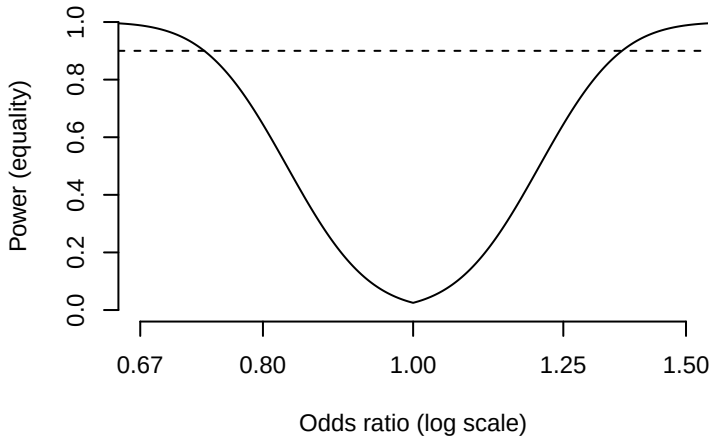

Figure 2: Two arm with 700 per arm.

Figure 3 presents power curves under various sample sizes assuming a two-sided size 0.05 test with no missingness.

Two arms with equal assignment of 700 participants each results in power 0.8 for a two-sided test of size $\alpha = 0.05$ assuming a common odds ratio of 1.31 in the investigational group versus control. Two arms with equal assignment of 1,150 participants each results in power 0.8 for a two-sided test of size $\alpha = 0.05$ assuming a common odds ratio of 1.24 in the investigational group versus control.

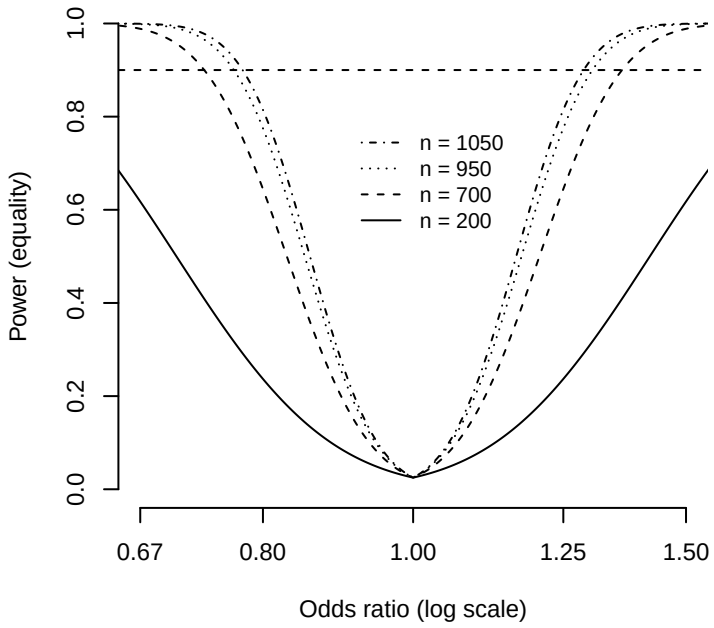

Figure 3: Power for two-arm comparison assuming fixed sample size (per arm as labelled) test of size $\alpha = 0.05$ for equality.

3 Simulations

3.1 Model

Let $Y_i \in \{1, \dots, 8\}$ be the day 14 clinical health status for participant i where 1 is best and 8 is worst (death). Let $\tau_i \in \{1, 2, 3\}$ denote treatment arm and x_i the corresponding design encoding such that $\eta_\tau = x_\tau^\mathsf{T}\beta$ is the shift in the log-odds of each outcome for participant receiving treatment τ. For the (reduced) primary model we specify:

$$P(Y_i \leq k|\alpha, \beta, x_i) = \text{logit}^{-1}(\alpha_k + x_i^T\beta) \quad i = 1, 2, ..., k = 1, ..., 8,$$

where $\{\alpha_k\}$ is increasing in k and $\beta = (\beta_1, \beta_2)^T$ represent the cumulative log-odds ratio of treatments 2 (candesartan) and 3 (candesartan + repagermanium) relative to control group 1.

These simulations assume treatment coding according to

$$X = \begin{array}{c} \begin{matrix} x_1 & x_2 \end{matrix} \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \text{placebo + placebo} \\ \text{candesartan + placebo} \\ \text{candesartan + repagermanium,} \end{array}$$

so arm 1 is P+P, arm 2 is C+P and arm 3 is C+R.

Define the outcome level probabilities in the control group by $\pi_k = P[Y = k|\alpha]$, for $k = 1, ..., 8$ where

$$\pi_k = \begin{cases} 1 - \text{logit}^{-1}(\alpha_1) & \text{if } k = 1 \\ \text{logit}^{-1}(\alpha_{k-1}) - \text{logit}^{-1}(\alpha_k) & \text{if } k \in \{2, ..., 7\} \\ \text{logit}^{-1}(\alpha_7) & \text{if } k = 8 \end{cases}$$

We specify the priors on the level probabilities as Dirichlet on the simplex:

$$\pi \sim \text{Dirichlet}(\kappa)$$

for some $\kappa = (\kappa_1, ..., \kappa_8)$, ideally informed from prior data where the sum $\sum \kappa_k$ gives the concentration. For the current simulations, $\kappa_k = 1/4$ for all k.

For the coefficients the current simulations assume the prior

$$\beta_1, \beta_2 \sim \text{Normal}(0, \sigma_\beta = 1).$$

If any of the outcome levels have zero observations across all treatment groups, then that outcome level is removed from the model for that analysis.

3.2 Decision Rules

In terms of the above model, we consider two possible decision processes.

3.2.1 Effectiveness

Effectiveness of the intervention against each control group is assessed by (conditional on data from n participants)

$$\mathbb{P}(\beta_3 + \beta_2 > \delta | \texttt{data}_n) \quad \text{(C+R > P+P)} \tag{1}$$

$$\mathbb{P}(\beta_3 > \delta | \texttt{data}_n) \quad \text{(C+R > C+P)} \tag{2}$$

$$\mathbb{P}(\beta_3 + \beta_2 > 0, \beta_3 > \delta | \texttt{data}_n) \quad \text{(C+R > P+P and C+P).} \tag{3}$$

If both probabilities (1) and (2) exceed some threshold, ϵ_{eff} then a decision of effectiveness relative to each control is recommended, and the trial may stop early. If either probability (1) or (2) falls below some threshold, ϵ_{hrm}, then a decision of harmfulness relative to at least one control is recommended, and the trial may stop early. The simulations specify $\delta = 0$.

Alternatively, decision could use the joint assessment (3).

3.2.2 Trial Futility

Alternatively, futility of the trial itself may be monitored according to the predicted probability of deciding effectiveness at the maximum sample size (assuming this can not be increased) conditional on data from n participants assuming an additional m participants will be recruited

$$\text{PPoS}_{\text{P+P}}(\texttt{data}_n, m) = \mathbb{E}[\mathbb{P}(\beta_3 + \beta_2 > \delta | \texttt{data}_{n+m}) | \texttt{data}_n] \tag{4}$$

$$\text{PPoS}_{\text{C+P}}(\texttt{data}_n, m) = \mathbb{E}[\mathbb{P}(\beta_3 > \delta | \texttt{data}_{n+m}) | \texttt{data}_n] \tag{5}$$

$$\text{PPoS}_{\text{P+P,C+P}}(\texttt{data}_n, m) = \mathbb{E}[\mathbb{P}(\beta_3 + \beta_2 > \delta, \beta_3 > 0 | \texttt{data}_{n+m}) | \texttt{data}_n] \tag{6}$$

In practice, calculation of these quantities requires assumptions about the future distribution of covariates included in the model. In the simulations, the only covariate is assigned treatment which is assumed to be assigned 1:1:1 up to the maximum sample size.

3.3 Scenarios

These simulations assess two decision processes:

- interim decisions based on posterior probability of effectiveness and harm only
- interim decisions based on posterior probability of effectiveness and predictive probability of success (trial futility)

The simulations apply the following decision rules:

- rule 1 (trial futility):
  - stop for effectiveness if both (1) and (2) exceed ϵ_{eff}
  - stop for futility if either (4) or (5) fall below ϵ_{fut}
- rule 2 (intervention harm):
  - stop for effectiveness if both (1) and (2) exceed ϵ_{eff}
  - stop for harm if either (1) or (2) falls below ϵ_{hrm}

The simulations consider the following thresholds for effectiveness: $\epsilon_{\text{eff}} \in \{0.95, 0.975\}$. The simulations assess a variety of thresholds for harm and futility, from 0 to 0.25.

The simulations investigate scenarios with the following effect sizes (odds ratios) on the investigational arm: (1/1.1, 1.0, 1.1, 1.2, 1.3, 1.5, 2.0). All scenarios currently assume equivalence of the two control arms.

Simulations are conducted for a three arm and a two arm trial. In the three arm scenario, the simulations do not currently consider arm dropping. All three arms are assigned participants until the trial stops.

3.4 Computational Details

At each interim analysis, posterior draws are generated using HMC via Stan using `rstan::sampling`. To calculate PPoS, a sub-sample is used to generate 500 posterior predictive draws as imputed future data, the posteriors approximated via Laplace's method using `rstan::optimizing`, and the required quantities are then derived.

For each posterior probability scenario, 1,000 trial replications are simulated. For each PPoS scenario, 500 trial replications are simulated.

4 Operating Characteristics

The following figures present:

- the Type I error of each decision rule
- the power for each decision rule
- the probability of stopping early for effectiveness, harm, or futility

In summary,

- type I error is about 0.05 for a decision rule which requires the pariwise odds ratio with each control to be greater than 1 with probability 0.93.
- under this decision rule, power is greater than 0.89 for an odds ratio for 1.3 over two equivalent control groups and 0.65 for an odds ratio of 1.2 over two equivalent control groups.

4.1 Posterior Probabilities

4.1.1 Three Arms

Table 1: Type I error for deciding investigational arm effective relative to both controls.

	ϵ_{hrm}										
ϵ_{eff}	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.2
0.90	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
0.91	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
0.92	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
0.93	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.94	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
0.95	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
0.96	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
0.97	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
0.98	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.99	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 2: Power for deciding investigational arm effective relative to both controls.

	ϵ_{eff}, $(\epsilon_{\text{hrm}} = 0.05)$									
OR	0.9	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
0.91	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00
1.00	0.07	0.07	0.06	0.05	0.04	0.04	0.03	0.02	0.01	0.01
1.10	0.38	0.35	0.31	0.29	0.25	0.22	0.18	0.14	0.09	0.06
1.20	0.72	0.69	0.67	0.65	0.61	0.55	0.51	0.44	0.35	0.23
1.30	0.93	0.91	0.90	0.89	0.86	0.84	0.81	0.77	0.69	0.57
1.50	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99	0.98	0.96
2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3: Power for deciding investigational arm harmful relative to at least one control.

	$\epsilon_{\text{hrm}}, (\epsilon_{\text{eff}} = 0.93)$										
OR	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.2
0.91	0.24	0.35	0.43	0.50	0.54	0.59	0.63	0.66	0.68	0.70	0.85
1.00	0.05	0.09	0.13	0.16	0.19	0.23	0.26	0.29	0.32	0.35	0.55
1.10	0.01	0.02	0.03	0.04	0.05	0.07	0.08	0.09	0.10	0.11	0.24
1.20	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.02	0.03	0.08
1.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
1.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

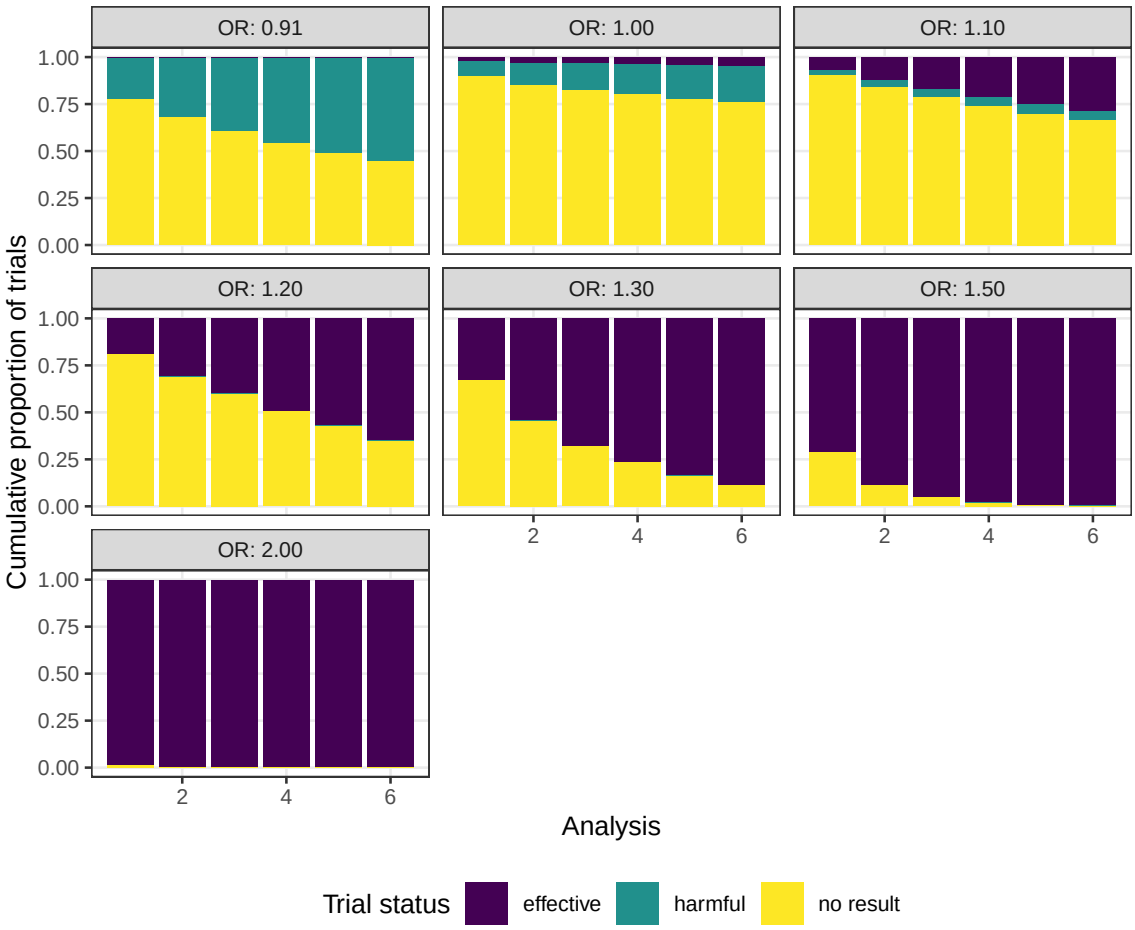

Figure 4: Cumulative trial status by analysis, $\epsilon_{\text{eff}} = 0.93$ and $\epsilon_{\text{hrm}} = 0.05$.

4.2 Predictive Probability of Success (PPoS)

Table 4: Probability of effectiveness decision as a function of odds ratio in investigational arm.

		Futility Threshold							
ϵ_{eff}	OR	0	0.02	0.025	0.05	0.1	0.15	0.2	0.25
0.950	0.91	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.950	1.00	0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.05
0.950	1.10	0.19	0.19	0.18	0.17	0.17	0.16	0.15	0.14
0.950	1.20	0.57	0.56	0.56	0.54	0.51	0.46	0.43	0.39
0.950	1.30	0.85	0.84	0.84	0.82	0.80	0.77	0.73	0.68
0.950	1.50	0.99	0.99	0.99	0.98	0.96	0.95	0.94	0.92
0.950	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.975	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.975	1.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
0.975	1.10	0.10	0.10	0.10	0.09	0.09	0.08	0.07	0.07
0.975	1.20	0.42	0.41	0.40	0.39	0.35	0.32	0.30	0.29
0.975	1.30	0.74	0.72	0.72	0.71	0.68	0.64	0.61	0.58
0.975	1.50	0.98	0.97	0.97	0.95	0.95	0.93	0.91	0.88
0.975	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 5: Power for futility decision as a function of odds ratio in investigational arm.

		Futility Threshold						
ϵ_{eff}	OR	0.02	0.025	0.05	0.1	0.15	0.2	0.25
0.950	0.91	0.98	0.98	0.98	0.99	0.99	0.99	0.99
0.950	1.00	0.87	0.88	0.90	0.93	0.94	0.94	0.95
0.950	1.10	0.64	0.67	0.73	0.77	0.80	0.83	0.85
0.950	1.20	0.26	0.28	0.34	0.42	0.50	0.55	0.60
0.950	1.30	0.08	0.09	0.12	0.17	0.22	0.26	0.32
0.950	1.50	0.01	0.01	0.02	0.04	0.05	0.06	0.08
0.950	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.975	0.91	0.99	0.99	0.99	1.00	1.00	1.00	1.00
0.975	1.00	0.91	0.92	0.95	0.96	0.98	0.98	0.98
0.975	1.10	0.76	0.77	0.83	0.86	0.89	0.90	0.92
0.975	1.20	0.37	0.40	0.48	0.56	0.62	0.67	0.69
0.975	1.30	0.15	0.16	0.20	0.26	0.33	0.37	0.40
0.975	1.50	0.01	0.02	0.04	0.05	0.07	0.09	0.11
0.975	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

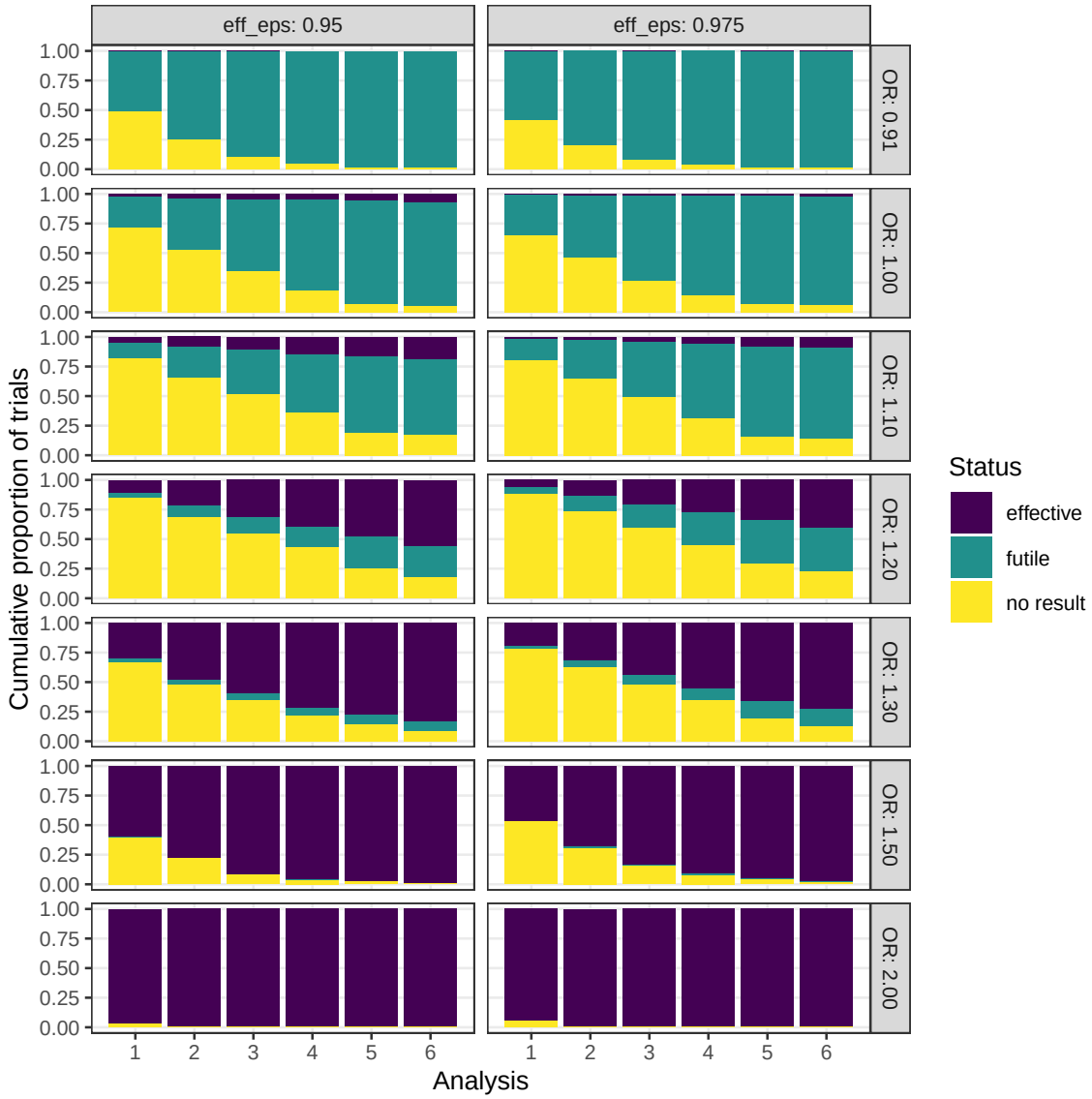

Figure 5: Cumulative status of simulated trials at each analysis, where $\epsilon_{\text{fut}} = 0.02$. To maintain proportions, trial status is carried forward after stopping, e.g. if a trial was futile at analysis 1, then it is also futile at analysis 2 to 6

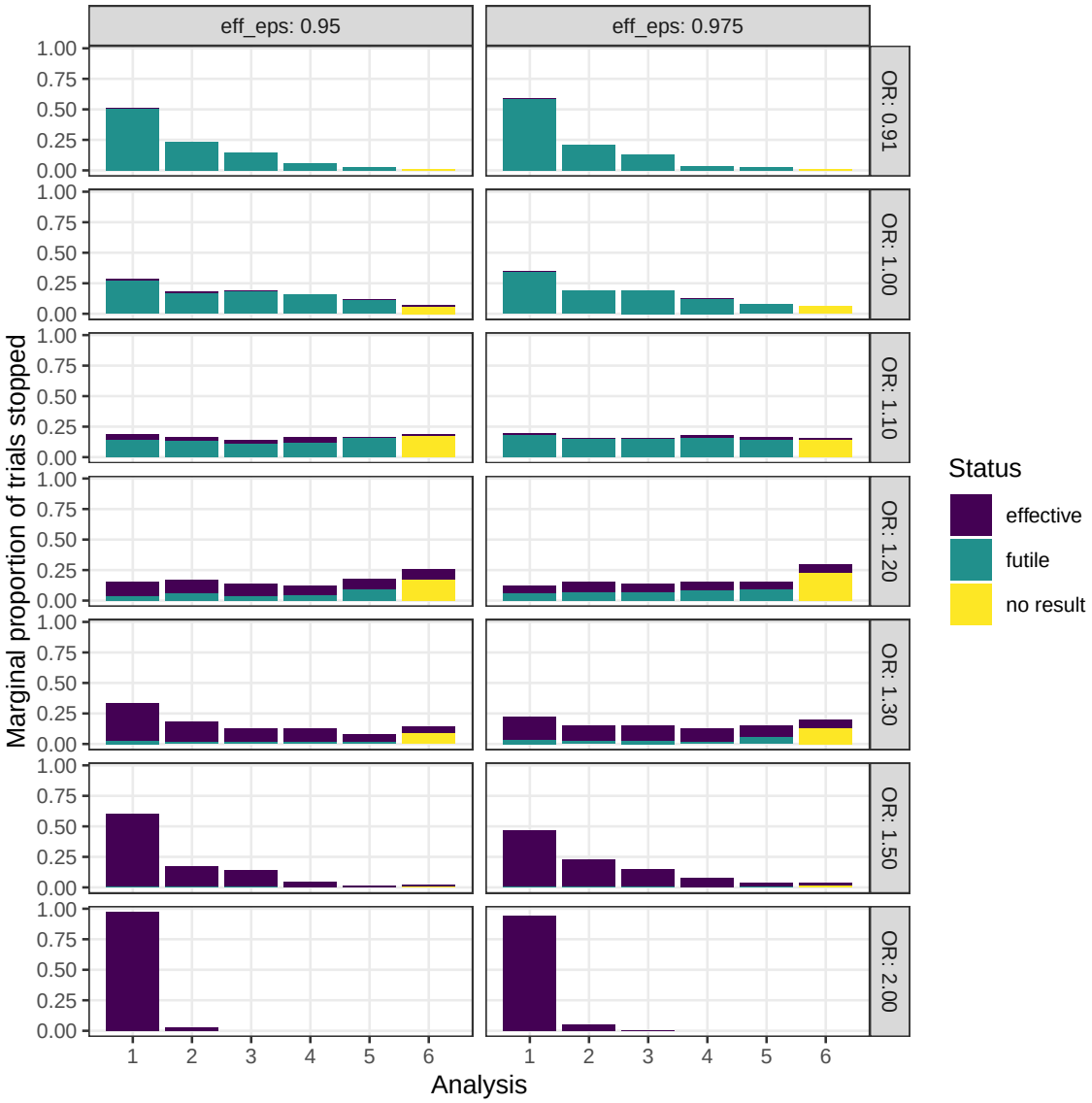

Figure 6: Marginal status of simulated trials by analysis, where $\epsilon_{\text{fut}} = 0.02$.

5 Appendix - Additional Summaries

5.1 Posterior

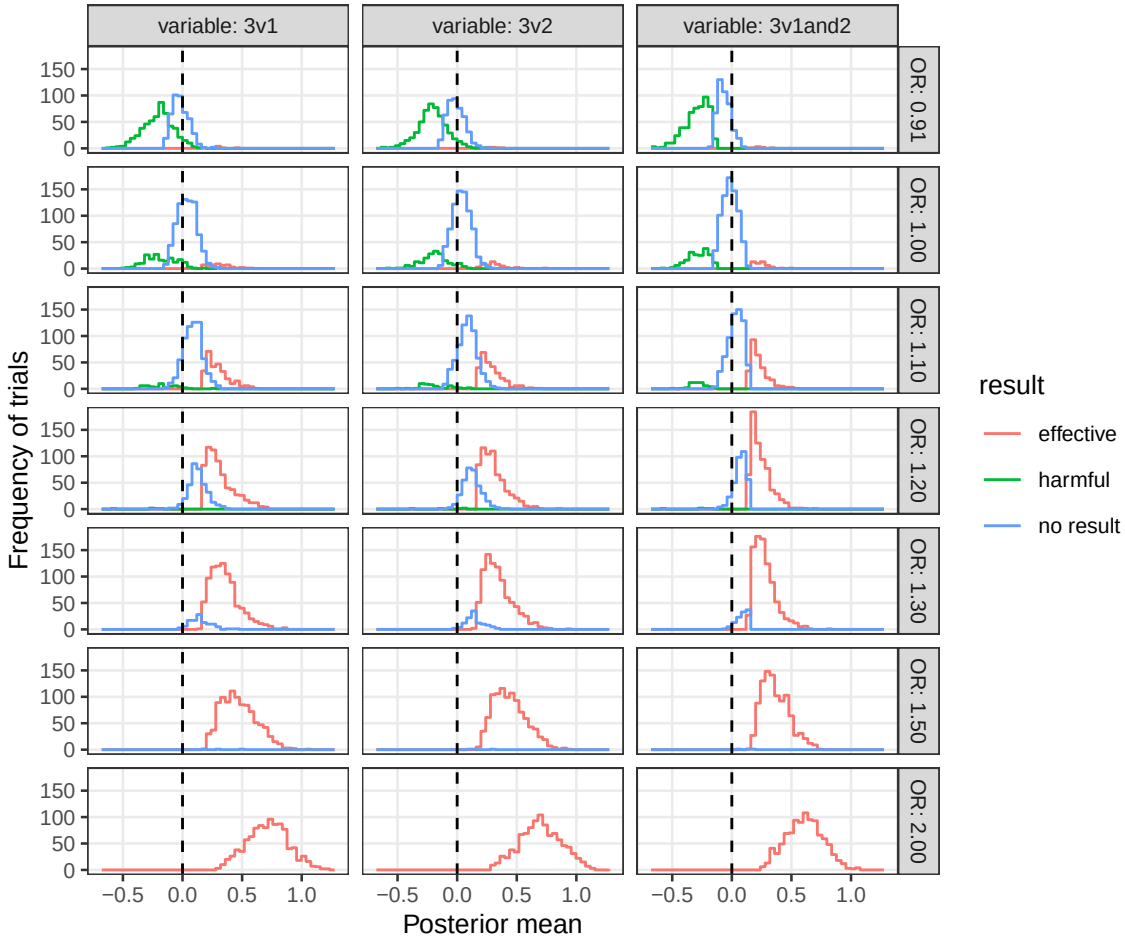

Figure 7: Posterior expectation at stopping.

5.2 Predictive Probability of Success

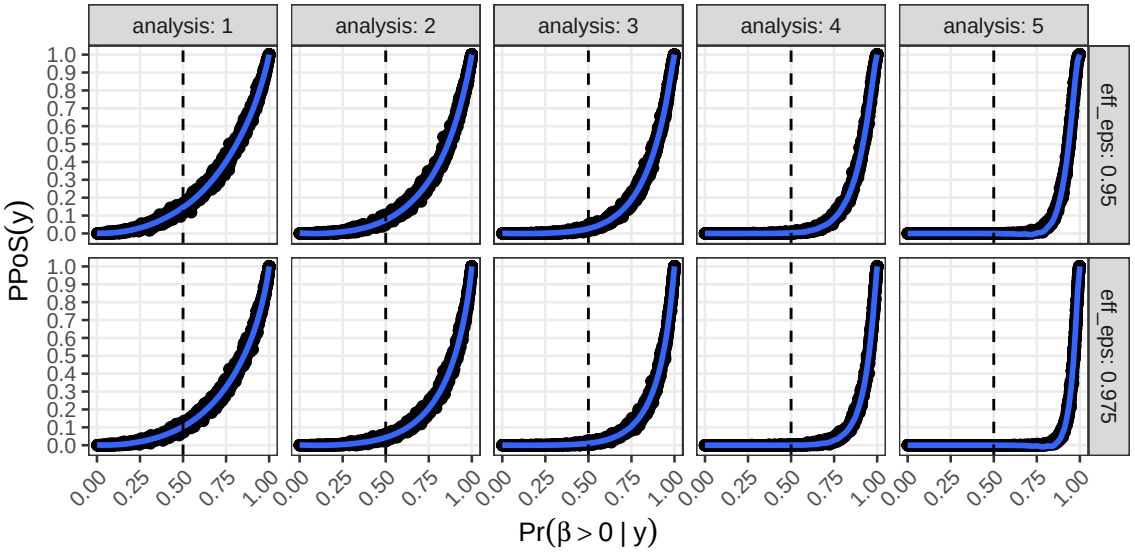

Figure 8: Estimated relationship of predicted probability of success (PPoS), posterior probability of effectiveness, and sample size for contrast of arm 3 vs arm 1.