Comparison of inferential Bayesian seamless phase II/III designs for acute stroke trials with biomarkers as surrogate endpoints

Marcio Augusto Diniz
Cedars-Sinai Medical Center

August 14, 2020

Joint work with Patrick Lyden and Mourad Tighiouart

Introduction

Drug development

Introduction

Drug development

- Traditionally, it follows a series of stages either in academia or industry before being evaluated by the U.S. Food and Drug Administration (FDA):

Introduction

Drug development

- Traditionally, it follows a series of stages either in academia or industry before being evaluated by the U.S. Food and Drug Administration (FDA):
  - Pre-clinical studies: Discovery;
  - Phase I: Safety;
  - Phase IIa/IIb: Activity/Efficacy;
  - Phase III: Definitive Evidence of Efficacy;
  - New Drug Application (NDA) for FDA;
  - Phase IV: Safety over time.

Introduction

Drug development

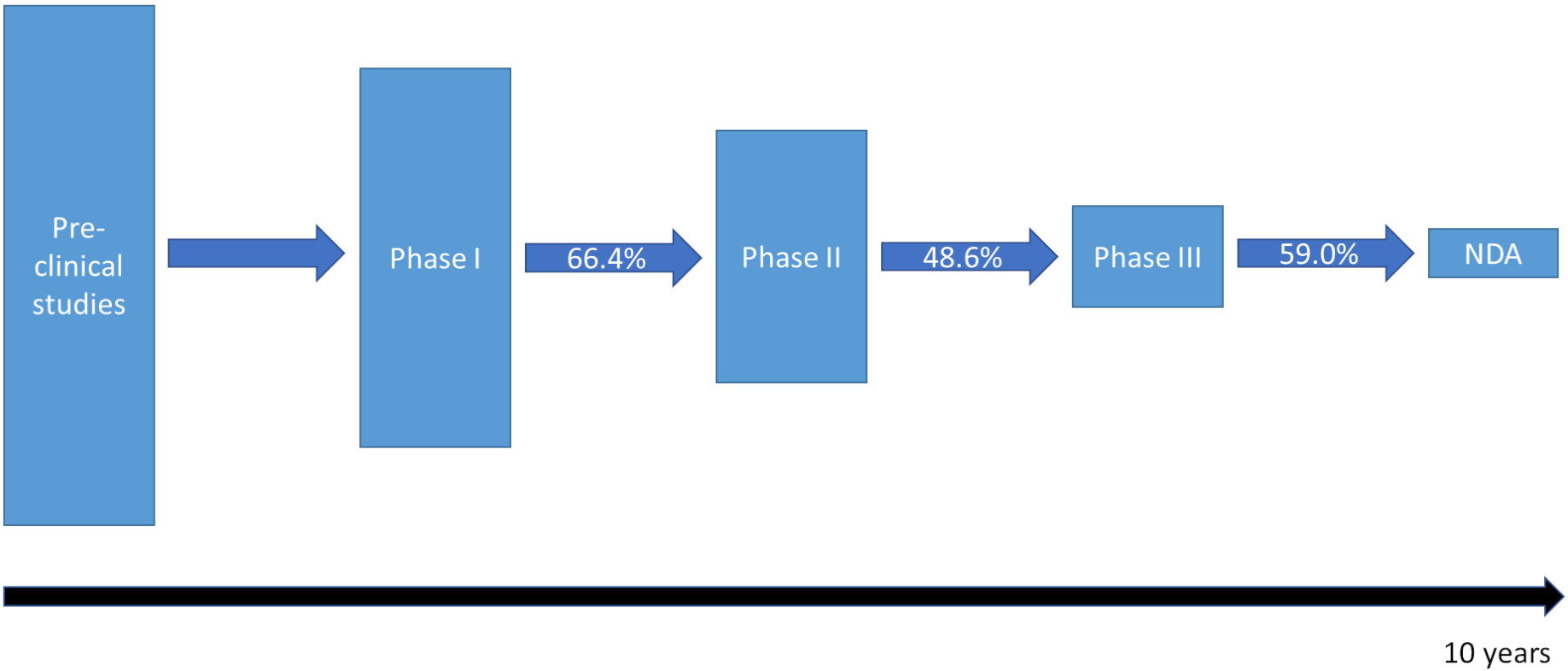


Figure: Estimates from Wong et al. (2019) based on 406 038 entries of clinical trial data for over 21 143 compounds from January 1, 2000 to October 31, 2015

Introduction

Drug development

Could we improve such process?

- Phase III trials require significant resources - time, money and patients;
- Overall failure rate of 41% for phase III trials, varying from 24.7% for infectious diseases to 74.5% in oncology Grayling et al. (2019);
- Several authors Vickers et al. (2007); Minnerup et al. (2014); Jardim et al. (2017) have pointed out that phase II trials are responsible for the high rates of negative phase III trials:
  - Design of single-arm instead of comparative randomized studies Taylor et al. (2006); Tang et al. (2010);
  - The use of short-term endpoints as a surrogate to long-term endpoints that will be used in phase III trials Stroke (2001); Wilson et al. (2015).

Introduction

Drug Development

Single-arm designs

- One-sample test;
- Disadvantages:
  - No accounting for sampling error in control estimates;
  - Differences in case-mix;
- Advantages:
  - Smaller samples sizes;
  - Shorter trial duration.

Randomized designs

- Two-sample test;
- Advantages:
  - Accounting for sampling error in control estimates;
  - Comparable case-mix;
- Disadvantages:
  - Larger samples sizes;
  - Longer trial durations;
  - Clinical Equipose.

Introduction

Drug development

Clinical Equipose

- It is the principle that states there is community uncertainty about the relative therapeutic merits across all arms;
- All patients enrolled in a trial can be assured of receiving nothing less than competent medical care. Hey and Kimmelman (2015)

Introduction

Drug development

Clinical Equipose

- It is the principle that states there is community uncertainty about the relative therapeutic merits across all arms;
- All patients enrolled in a trial can be assured of receiving nothing less than competent medical care. Hey and Kimmelman (2015)

Balanced Randomization

- It might not be appealing to patients know that they might not be enrolled in the experimental arm;
- Ethical dilemma when subjects are equally randomized clashing with patient's best interest and clinical practice.

Introduction

Drug development

Response Adaptive Randomization (RAR)

- It has been proposed Thompson (1933); Wei and Durham (1978); Eisele (1994); Berry and Eick (1995); Ivanova (2003) under classical and Bayesian paradigms;
- On average, patients are allocated to the most promising experimental arms;
- Controversial for two-arm studies Hey and Kimmelman (2015); Korn and Freidlin (2011); Thall et al. (2015);
- Although it is an useful strategy in the context of dose selection (multi-arm studies) Meinzer et al. (2017).

Introduction

Drug development

Response Adaptive Randomization (RAR)

- It has been proposed Thompson (1933); Wei and Durham (1978); Eisele (1994); Berry and Eick (1995); Ivanova (2003) under classical and Bayesian paradigms;
- On average, patients are allocated to the most promising experimental arms;
- Controversial for two-arm studies Hey and Kimmelman (2015); Korn and Freidlin (2011); Thall et al. (2015);
- Although it is an useful strategy in the context of dose selection (multi-arm studies) Meinzer et al. (2017).

Randomized phase II trials with RAR

- It requires larger samples sizes than the single-arm studies;
- Changes on the paradigm are limited by the availability of resources.

Introduction

Drug development

Inferential seamless phase II/III

- It has been proposed in the literature Maca et al. (2006); Bretz et al. (2006) to shorten the drug development process with the gap between phase II and III being minimized and make efficiently use of patients' data;
- In the first stage, a randomized phase II trial is performed such that active arm is selected comparing to the control arm based on a short-term endpoint;
- In the second stage, a phase III trial is implemented with the long-term endpoint such that the data from the patients of first stage is also taken into account.
- Such framework allows to accommodate more sophisticated phase II designs;
- It allow us to take into account type I error in both phases.

Introduction

Drug development

Inferential seamless phase II/III

- Inoue et al. (2002) proposed a seamless design under the Bayesian approach with a joint Bayesian model for a short-term multinomial and a time-to-event endpoints such that future event times were simulated given the current data at each interim analysis;
- Huang et al. (2009) introduced RAR in a phase II/III design while also jointly modeled a time-to-event and a multinomial endpoint under a Bayesian approach;
- Others have also proposed similar designs under a hybrid Bayesian/classical and classical approaches.

Introduction

Stroke

- Acute stroke is a sudden interruption in the blood supply of the brain, injuring brain cells and tissues.

Modified Rankin Scale (mRS)

- It is a 7-level scale proposed by John Rankin in 1957:
  - 0: No symptoms;
  - 1: No significant disability. Able to carry out all usual activities, despite some symptoms;
  - 2: Slight disability. Able to look after own affairs without assistance, but unable to carry out all previous activities;
  - 3: Moderate disability. Requires some help, but able to walk unassisted;
  - 4: Moderately severe disability. Unable to attend to own bodily needs without assistance, and unable to walk unassisted;
  - 5: Severe disability. Requires constant nursing care and attention, bedridden, incontinent;
  - 6: Dead.

Introduction

Stroke

Trials

- Stroke trials commonly have 90-day mRS as primary endpoint;
- mRS is often dichotomized as 0-1 or 0-2.

NIHSS

- National Institute of Health Stroke Scale (NIHSS) is neurological function measure ranging from 0= no deficit to 42 = extreme deficit;
- It is often assessed at baseline, 24-hours, 7-day and 90-day;
- 24h-NIHSS has sensitivity = 83% and specificity 81% based on IMA and IMS-II trials;
- Nowacki et al. (2017) proposed to use NIHSS as a surrogate of mRS in the adaptive randomization under the classical approach;

Introduction

Research questions

- What is the performance when we use NIHSS as a surrogate of mRS with the RAR under the Bayesian approach?
- Can we also use NIHSS to take decisions in addition to the RAR algorithm?

Contents

Modeling - Phase II

Long-term endpoint model

- L_{ij} be a binary indicator of 90-day mRS 0-2 that will be observed after a period τ_L for patient i receiving treatment j at time T_{ij}:
  - n_j patients were accrued for treatment j;
  - m_j patients were accrued for treatment j, but did not have their long-term endpoint observed at time t_{m_j};
  - $i = 1, \ldots, (n_j - m_j)$;
  - $j = 0, \ldots, J$, with $j = 0$ indicating the control arm.
- $L_{ij} \sim Bernoulli(\theta_j)$;
- θ_j is the probability of the event of interest for the long-term endpoint.

Modeling - Phase II

Long-term endpoint model

Likelihood

$$L(\theta_j | D_{n_j}(t_{m_j})) = \prod_{i=1}^{n_j - m_j} \theta_j^{l_{ij}} (1 - \theta_j)^{1 - l_{ij}},$$

for $j = 1, \ldots, J$.

Prior distributions

$$\theta_j \sim beta(a_j, b_j),$$

for $j = 1, \ldots, J$.

Modeling - Phase II

Long-term endpoint model

Posterior distribution

$$\theta_j | D_{n_j}(t_{m_j}) \sim beta\left(a_j + \sum_{i=1}^{n_j - m_j} l_{ij}, b_j + (n_j - m_j) - \sum_{i=1}^{n_j - m_j} l_{ij}\right),$$

for $j = 1, \ldots, J$.

Modeling - Phase II

Long-term endpoint model

Issues

- Depending on accrual rate of patients, m_j will be greater than zero;
- Often a short-term endpoint S_{ij} for patients $i = n_j - m_j + 1, \ldots, n_j$ is available;
- Which strategy can we adopt in our clinical trial?
  - Draw inferences for $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_J)$ based only on the patients that the long-term endpoint is observed;
  - Replace the long-term endpoint by the short-term endpoint when the former is not available in the likelihood of the long-term model.

Modeling - Phase II

Short-term endpoint model

- S_{ij} be a binary indicator of the NIHSS ≤ 10 that will be observed after a period τ_S with $\tau_S < \tau_L$ for patient i receiving treatment j at time T_{ij}:
  - n_j patients were accrued for treatment j;
  - m_j patients were accrued for treatment j, but did not have their long-term endpoint observed at time t_{m_j};
  - $i = 1, \ldots, (n_j - m_j)$;
  - $j = 0, \ldots, J$, with $j = 0$ indicating the control arm.
- $S_{ij}|L_{ij} = l \sim Bernoulli(\lambda_l)$;
- $\lambda_l = P(S_j = 1|L_j = l)$ such that $1 - \lambda_0$, λ_1 are the bio-marker sensitivity and specificity.

Modeling - Phase II

Short-term endpoint model

Likelihood

$$L(\theta_j, \lambda_1, \lambda_0 | D_{n_j}(t_{m_j})) = \prod_{i=1}^{n_j - m_j} \theta_j^{l_{ij}} (1-\theta_j)^{1-l_{ij}} \times \prod_{i=n_j-m_j+1}^{n_j} \{(\theta_j \lambda_1 + (1-\theta_j)\lambda_0)^{s_{ij}} \times (\theta_j[1-\lambda_1] + [1-\theta_j][1-\lambda_0])^{1-s_{ij}}\},$$

for $j = 1, \ldots, J$.

Modeling - Phase II

Short-term endpoint model

Issues

- In the case the short-term endpoint is a perfect bio-marker,
  - Then $\lambda_1 = P(S_j = 1|L_j = 1) = 1$ and $\lambda_0 = P(S_j = 1|L_j = 0) = 0$;
  - The likelihood reduces to

$$L(\theta_j, \lambda_1, \lambda_0 | D_{n_j}(t_{m_j})) = \prod_{i=1}^{n_j - m_j} \theta_j^{l_{ij}} (1-\theta_j)^{1-l_{ij}} \times \prod_{i=n_j-m_j+1}^{n_j} \theta_j^{s_{ij}} (1-\theta_j)^{1-s_{ij}};$$

- Otherwise, posterior estimates for $\boldsymbol{\theta}$ will be biased due confounding with λ_1 and λ_0.

Modeling - Phase II

Long-term and short-term endpoints joint model

- Following Daniel Paulino et al. (2003), let R_{ijls} be a binary indicator for patient i receiving treatment j with $L_{ij} = l$ and $S_{ij} = s$ at time T_{ij}:
  - n_j patients were accrued for treatment j;
  - m_j patients were accrued for treatment j, but did not have their long-term endpoint observed at time t_{m_j};
  - $i = 1, \ldots, (n_j - m_j)$;
  - $j = 0, \ldots, J$, with $j = 0$ indicating the control arm;
  - $l, s = 0, 1$.

Modeling - Phase II

Long-term and short-term endpoints joint model

- $\mathbf{R}_{ij} = (R_{ij00}, R_{ij10}, R_{ij01}, R_{ij11}) \sim multinomial(1, \mathbf{p})$ with $\mathbf{p}_j = (p_{j00}, p_{j01}, p_{j10}, p_{j11})$ where
  - $p_{j00} = P(L_{ij} = 0 \text{ and } S_{ij} = 0) = [1 - \theta_j][1 - \lambda_0]$;
  - $p_{j01} = P(L_{ij} = 0 \text{ and } S_{ij} = 1) = [1 - \theta_j]\lambda_0$;
  - $p_{j10} = P(L_{ij} = 1 \text{ and } S_{ij} = 0) = \theta_j[1 - \lambda_1]$;
  - $p_{j11} = P(L_{ij} = 1 \text{ and } S_{ij} = 1) = \theta_j\lambda_1$.

Modeling - Phase II

Long-term and short-term endpoints joint model

Likelihood

- I_j is an index set of patients that have data for both endpoints in arm j;
- After $|I_j|$ accrued patients,
  $D_{I_j}(t_{n_j-|I_j|}) = \{(s_{ij}, l_{ij}, t_{ij}) : t_{ij} + \tau_L < t_I \text{ for } i \in I_j\}$
- $D_I(t_{n-|I_j|}) = \cup_{j=0}^{J} D_{I_j}(t_{n_j-|I|})$.

for $j = 1, \ldots, J$.

Modeling - Phase II

Long-term and short-term endpoints joint model

Likelihood

$$L(\boldsymbol{\theta}, \lambda_1, \lambda_0 | D_I(t_{n-|I|})) = \prod_{j=1}^{J} \prod_{i \in I} p_{j00}^{r_{ij00}} p_{j01}^{r_{ij01}} p_{j10}^{r_{ij10}} p_{j11}^{r_{ij11}}$$

$$= \prod_{j=1}^{J} \prod_{i \in I} \theta_j^{r_{ij11}+r_{ij10}} (1-\theta_j)^{(r_{ij00}+r_{ij01})} \times$$

$$\lambda_1^{r_{ij11}} (1-\lambda_1)^{r_{ij10}} \times (1-\lambda_0)^{r_{ij00}} \lambda_0^{r_{ij01}}.$$

for $j = 1, \ldots, J$.

Modeling - Phase II

Long-term and short-term endpoints joint model

Prior distributions

$$\begin{aligned} \theta_j &\sim beta(a_j, b_j), \\ \lambda_l &\sim beta(c_l, c_l). \end{aligned}$$

for $j = 1, \ldots, J$ and $l = 0, 1$.

Modeling - Phase II

Long-term and short-term endpoints joint model

Posterior distribution

$$\begin{aligned}
\theta_j &\sim beta(a_j + \sum_{i \in I}(r_{ij11} + r_{ij10}), b_j + \sum_{i \in I}(r_{ij00} + r_{ij01})), \\
\lambda_1 &\sim beta(c_{j1} + \sum_{j=1}^{J}\sum_{i \in I} r_{ij11}, d_{j1} + \sum_{j=1}^{J}\sum_{i \in I} r_{ij10}), \\
\lambda_0 &\sim beta(c_{j0} + \sum_{j=1}^{J}\sum_{i \in I} r_{ij01}, d_{j0} + \sum_{j=1}^{J}\sum_{i \in I} r_{ij00}).
\end{aligned}$$

where $I = \cup_{j=1}^{J} I_j$ and $j = 1, \ldots, J$.

Modeling - Phase II

Long-term and short-term endpoints joint model

Posterior distribution

- $I_j = \{1, \ldots, n_j - m_j\}$ when we have accrued n_j patients for treatment j, but the long-term endpoint is missing for the last m_j patients.
- However, we also can augment our data such that the last m_j observations for the long-term endpoint will be generated from predictive distributions.

Modeling - Phase II

Long-term and short-term endpoints joint model

Predictive distribution

$$
\begin{aligned}
q_0 &= P(L_{ij} = 0 | S_{ij} = 0) \\
&= \frac{(1-\theta_j)(1-\lambda_0)}{(1-\theta_j)(1-\lambda_0) + \theta_j(1-\lambda_1)}, \\
q_1 &= P(L_{ij} = 1 | S_{ij} = 1) \\
&= \frac{\theta_j \lambda_1}{\theta_j \lambda_1 + (1-\theta_j)\lambda_0}.
\end{aligned}
$$

where for $i = n_j - m_j + 1, \ldots, n_j$ and $j = 1, \ldots, J$.

- q_0 and q_1 can be interpreted as Negative Predictive Value and Positive Predictive Value, respectively.

Modeling - Phase II

Long-term and short-term endpoints joint model

Predictive distribution using sensitivity

$$\sum_{i=n_j-m_j+1}^{n_j} r_{ij00} | D_{n_j,m_j} \sim Bin\left(m_j - \sum_{i=n_j-m_j+1}^{n_j} s_{ij}, q_0\right),$$

$$\sum_{i=n_j-m_j+1}^{n_j} r_{ij01} = m_j - \sum_{i=n_j-m_j+1}^{n_j} s_{ij} - \sum_{i=n_j-m_j+1}^{n_j} r_{ij00}.$$

Modeling - Phase II

Long-term and short-term endpoints joint model

Predictive distribution using specificity

$$\sum_{i=n_j-m_j+1}^{n_j} r_{ij11} | D_{n_j,m_j} \sim Bin\left(\sum_{i=n_j-m_j+1}^{n_j} s_{ij}, q_1 \right),$$

$$\sum_{i=n_j-m_j+1}^{n_j} r_{ij10} = \sum_{i=n_j-m_j+1}^{n_j} s_{ij} - \sum_{i=n_j-m_j+1}^{n_j} r_{ij11}.$$

Modeling - Phase II

Long-term and short-term endpoints joint model

Augmentation algorithm

- From this setup, we are able to draw inferences for θ_j for $j = 1, \ldots, J$ as follows:
  1. Choose adequate initial values for $\theta_j^{(0)}$, $\lambda_0^{(0)}$, $\lambda_1^{(0)}$ with $I = \{1, \ldots, n_j - m_j\}$;
  2. For $k = 1, \ldots, K$,
     a. Imputation step:
     Sample $r_{ij}^{(k)} = (r_{ij11}^{(k)}, r_{ij10}^{(k)}, r_{ij00}^{(k)}, r_{ij01}^{(k)})$ from the predictive distributions for $i = n_j - m_j + 1, \ldots, n_j$ given $\theta_j^{(k-1)}$, $\lambda_0^{(k-1)}$, $\lambda_1^{(k-1)}$;
     b. Posterior step:
     Sample $\theta_j^{(k)}$, $\lambda_0^{(k)}$, $\lambda_1^{(k)}$ from posterior distribution with $I = \{1, \ldots, n_j\}$ given $r_j^{(k)}$.

Modeling - Phase III

Long-term model

- Let Y_{ij} be the ordinal mRS scale that is observed after a time window τ_L for patient i receiving treatment arm j
- N_j patients were accrued for treatment j until time t_{m_j};
- m_j patients were accrued for treatment j, but did not have their long-term endpoint observed at time t_{m_j};
- $i = 1, \ldots, N_j - m_j$;
- $j = 0, 1$ with $j = 0$ indicating the control arm;
- $Y_{ij} \sim multinomial(g_j)$ with $\boldsymbol{g_j} = (g_{j1}, \ldots, g_{jK})$.

Modeling - Phase III

Long-term model

Likelihood

$$L(\boldsymbol{g_j}|D_{N_j}(t_{m_j})) = \prod_{i=1}^{N_j}\prod_{k=1}^{K} g_{jk}^{I(y_{ij}=k)},$$

for $j = 0, 1$.

Modeling - Phase III

Long-term model

Prior distribution

$$\boldsymbol{g_j} \sim Dir(\alpha_{j1}, \ldots, \alpha_{jK});$$

Posterior distribution

$$\boldsymbol{g_j}|D_{N_j}(t_{m_j}) \sim Dir\left(\alpha_{j1} + \sum_{i=1}^{N_j - m_j} I(y_{ij} = 1), \ldots, \alpha_{jK} + \sum_{i=1}^{N_j - m_j} I(y_{ij} = K)\right).$$

Modeling - Phase III

Long-term model

Weighted average

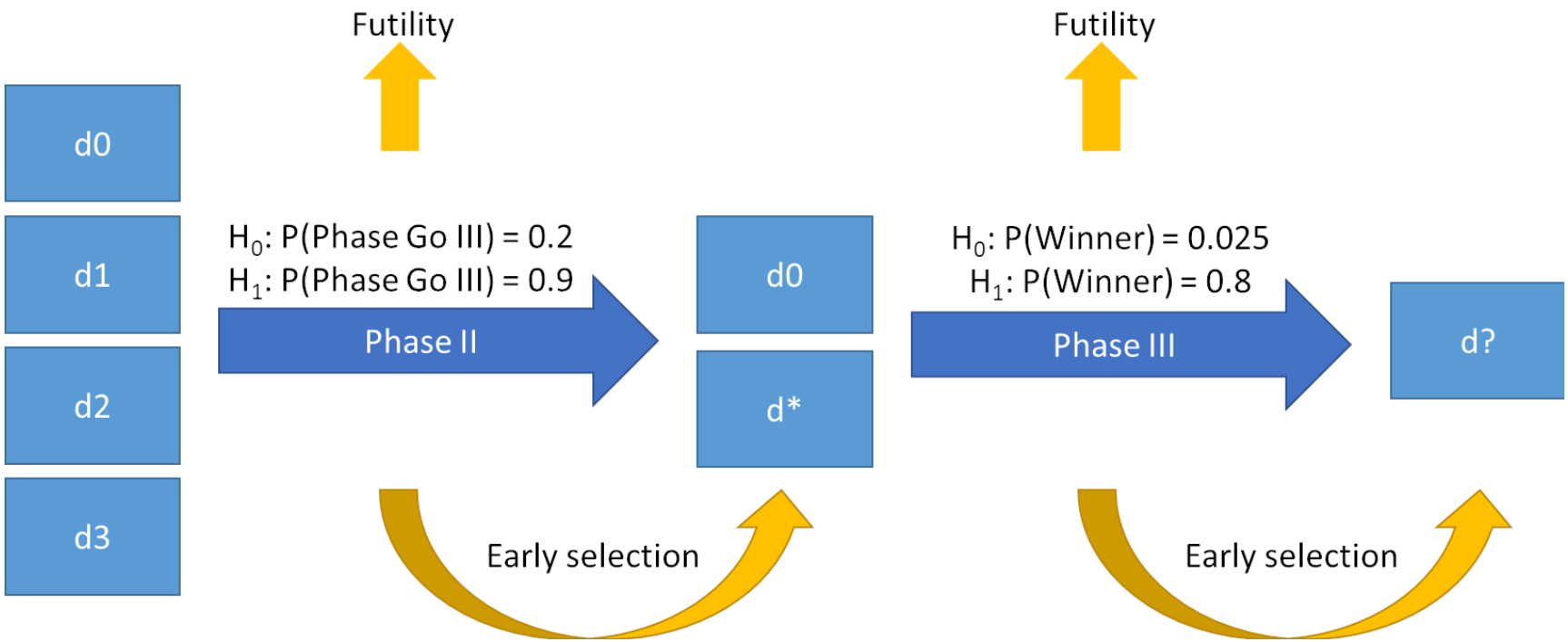
- We also assume that each category k of Y has an associated weight w_k;
- We are interested in the weighted average of the parameter vector $\boldsymbol{g}$,

$$\mu_j = \sum_{k=1}^{K} w_k g_{jk}.$$

- The posterior distribution of $\boldsymbol{\mu} = (\mu_0, \mu_1)$ is not analytically tractable, but it can estimated through simulations of the empirical distribution of $\boldsymbol{g}$.

Contents

Design - Seamless Phase II/III



Design - Phase II

Set up

- Goal: Select the lowest dose with 90% of efficacy relative to the maximum efficacy among J doses $(d_1, \ldots, d_J)$ compared to the control arm (d_0);
- Primary endpoint: Proportion of subjects who show 90-day mRS ≤ 2 or 7-day NIHSS ≤ 10;
- Efficacy: Higher proportion of events when compared to the control arm.

Randomization

- Stage 1: Balanced randomization of the first 15 patients for each of the $(J + 1)$ arms.
- Stake k: Adaptive Randomization every 5 patients.

Design - Phase II

RAR

Allocation probability

$$
\begin{aligned}
p_j &= P(\theta_j > max(\theta_{i\neq j})_{i=1,\ldots,J} | D_{n,m}) \\
&= \int_0^1 \ldots \int_{max(\theta_{i\neq j})_{i=1,\ldots,J}}^1 \pi_{\theta_j}(x_j | D_{n_j,m_j}) \prod_{i=1,i\neq j}^{J} \pi_{\theta_i}(x_i | D_{n_i,m_i}) \delta x_j \delta x_i,
\end{aligned}
$$

where $D_{n,m} = \cup_{j=0}^{J} D_{n_j}(t_{m_j})$ for $j = 1, \ldots, J$.

Design - Phase II

RAR

Allocation probability

$$P(\text{allocation arm } j | D_{n,m}) = \frac{p_j}{\sum_{j=0}^{J} p_j},$$

where

$$p_0 = P(\text{allocation arm } 0 | D_{n,m}) = \frac{1}{(J+1)}.$$

Design - Phase II

Stopping rules

Winner probability

$$\begin{aligned} P(\text{winner arm } j|D_{n,m}) &= P(\theta_j > \theta_0|D_{n,m}) \\ &= \int_0^1 \int_{\theta_0}^1 \pi_{\theta_j}(x_j|D_{n_j,m_j})\pi_{\theta_0}(x_0|D_{n_0,m_0})\delta x_j \delta x_0. \end{aligned}$$

for $j = 1, \ldots, J$.

Design - Phase II

Stopping rules

Early Loser

- If

$$P(\text{winner arm } j | D_{n,m}) < \delta_{EL},$$

and arm j has at least 30 patients, then

$$P(\text{allocation arm } j | D_{n,m}) = 0$$

until the next allocation probability update.

Design - Phase II

Stopping rules

Early Winner

- If

$$P(\text{winner arm } j | D_{n,m}) > \delta_{EW},$$

and arm j has at least 50 patients, then arm j is declared the early winner and the trial is stopped early, and the trial proceeds to its phase III.

Design - Phase II

Stopping rules

Futility

- If

$$P(\theta_j > \theta_{min} | D_{n,m}) < \delta_F,$$

where θ_{min} is fixed by the clinician, then arm j is declared futile and it is dropped until the end of the trial.

Design - Phase II

Stopping rules

Late Winner

- After all patients have been evaluated, if

$$P(\text{winner arm } j | D_{n,m}) > \delta_{LW},$$

then arm j is declared the winner and the trial proceeds to its phase III. Otherwise, no dose is selected and the trial is stopped.

Design - Phase II

Stopping rules

Efficacy 90%

- If there is more than one arm as winner, then the probability of 90% efficacy is calculated for the winners,

$$\begin{aligned} P(\text{winner arm } j|D_{n,m}) &= P(\theta_j > 0.9\theta_{max}|D_{n,m}) \\ &= \int_0^1 \int_{0.9\theta_{mx}}^1 \pi_{\theta_j}(x_j|D_{n_j,m_j}) \times \\ &\quad \pi_{\theta_{max}}(x_{max}|D_{n_{max},m_{max}})\delta x_j \delta x_{max}. \end{aligned}$$

where $\theta_{max} = \theta_j^*$ for $j^* = \arg\max_j p_j$.

Design - Phase III

Set up

- Goal: Compare the selected dose (d_{j*}) in phase II with the control arm (d_0);
- Two co-primary endpoints: proportion of mRS ≤ 2 from phase II trial and UW-mRS;
- Efficacy: Higher proportion of events and weighted mean when compared to the control arm.

Randomization

- Patients will be randomized to control and treatment arms according to an unbalanced allocation ratio in such a way that the expected number of patients in each arm equalizes at the end of the trial.

Design - Phase II

Stopping rules

Efficacy

- $H_0 : H_{01} \cap H_{02}$ vs $H_1 : H_{11} \cap H_{12}$:
  - $H_{01} : \theta_0 \geq \theta_{j*}$ and $H_{02} : \mu_0 \geq \mu_{j*}$;
  - $H_{11} : \theta_0 < \theta_{j*}$ and $H_{12} : \mu_0 < \mu_{j*}$;
- The alternative hypotheses is accepted if

$$P(\theta_{j*} > \theta_0 | D_n(t)) > \eta_.$$

and

$$P(\mu_{j*} > \mu_0 | D_n(t)) > \gamma_..$$

with

- $\eta_. = \eta_{EW}$ and $\gamma_. = \gamma_{EW}$ for the interim analyses;
- $\eta_. = \eta_{LW}$ and $\gamma_. = \gamma_{LW}$ for the final analysis.

Design - Phase II

Stopping rules

Futility

- $H_0 : H_{01} \cap H_{02}$ vs $H_1 : H_{11} \cap H_{12}$:
  - $H_{01} : \theta_0 \geq \theta_{j*}$ and $H_{02} : \mu_0 \geq \mu_{j*}$;
  - $H_{11} : \theta_0 < \theta_{j*}$ and $H_{12} : \mu_0 < \mu_{j*}$;
- The null hypotheses is accepted if

$$Pred(P(\theta_{j*} > \theta_0 | D_n(t)) > \eta_{LW} | D_n(t)) > \eta_F$$

and

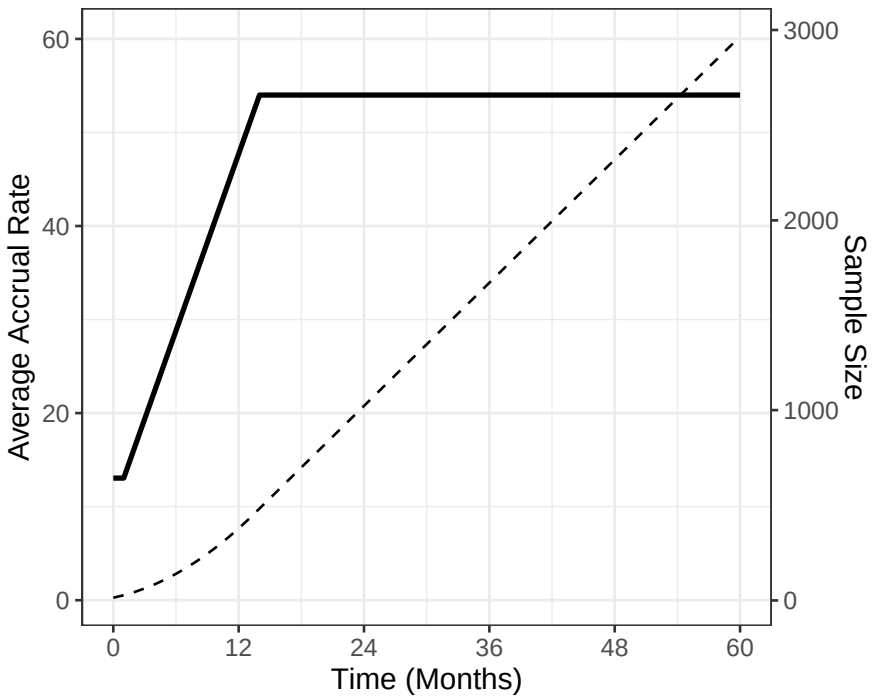
$$Pred(P(\mu_{j*} > \mu_0 | D_n(t)) > \gamma_{LW} | D_n(t)) > \gamma_F$$

where $Pred(.|D_n(t))$ indicates the predictive distribution.

Contents

Stroke trial

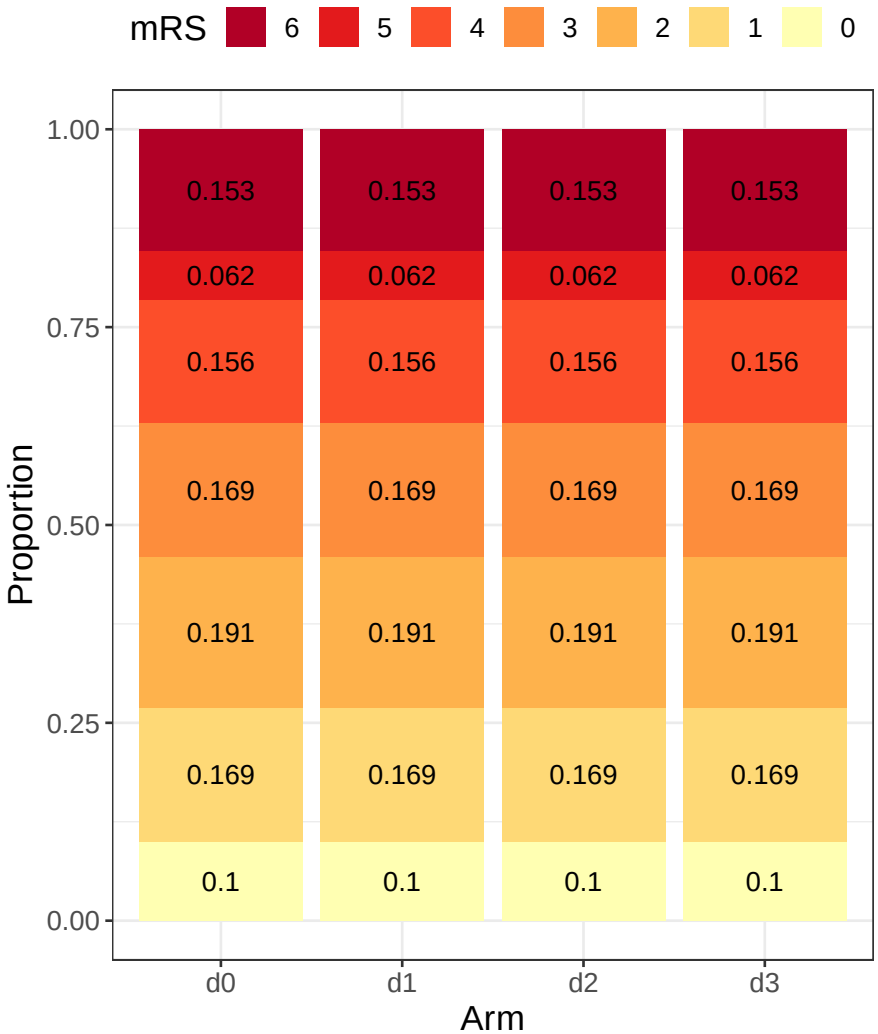
Accrual



- 20 European sites will all be ready to enroll on day 1;
- 100 Australian and US sites will ramp up to a total of 9 patients/month for 14 months;
- 0.45 patient/month/site;
- Total sample size: up to 3000.

Stroke trial

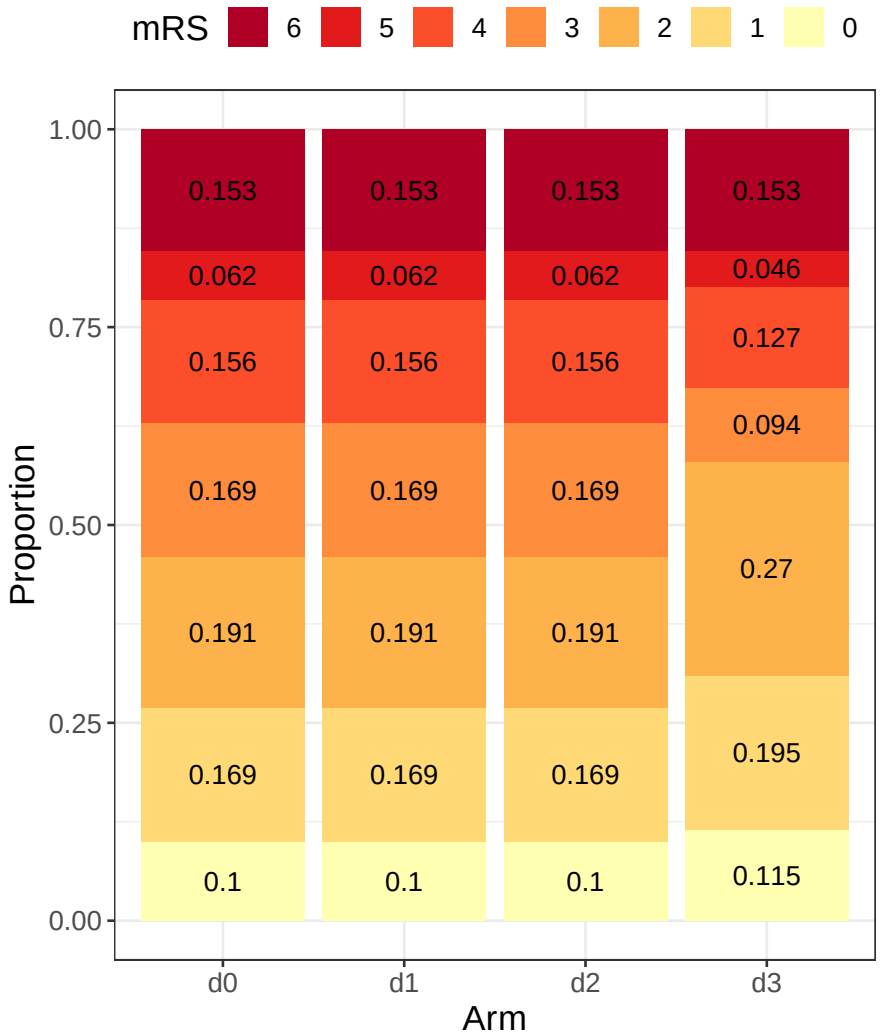
Null scenario



- For all arms,
  - mRS 0-2 = 0.46;
  - UW-mRS = 0.56028.

Stroke trial

Alternative scenario



- For arms d_0, d_1 and d_2,
  - proportion of mRS 0-2 = 0.46;
  - UW-mRS = 0.56028.
- For arm d_3,
  - proportion of mRS 0-2 = 0.58;
  - UW-mRS = 0.60066.

Stroke trial

Strategies

- S_0 : Long-term model;
- S_1 : Short-term model in the randomization and taking decisions;
- S_2 : Short term model only in the randomization;
- S_3 : Short and long-term joint model in the randomization and taking decisions.

Stroke trial

Strategy S_0

Prior Parameters

- Phase II: $(a_j, b_j) = (0.5, 0.5)$ for $j = 0, \ldots, 3$ were chosen as Jeffrey priors;
- Phase III: $(\alpha_{j0}, \ldots, \alpha_{j6})$ for $j = 0, \ldots, 3$ are defined based on data collected on phase II.

Design Parameters

- Parameters were adjusted based on 2000 simulated trials to satisfy type I error ≤ 0.025 and power > 0.80;
- Phase II:
  $(\delta_{EL} = 0.05, \delta_{EW} = 0.99, \delta_F = 0.10, \theta_{min} = 0.40, \delta_{LW} = 0.9)$;
- Phase III: $(\eta_{EW} = \gamma_{EW} = 0.97, \eta_{LW} = \gamma_{LW} = 0.95, \eta_F = \gamma_F = 0.05)$;

Stroke trial

Strategy S_0

Sample size

- Phase II = 800 and Phase III = 2100;

Operating Characteristics

- Null scenario:
  - Transition = P(Go to Phase III|Null scenario) = 0.192;
  - Type I error = P(Any winner in Phase III|Null scenario) = 0.0229;
  - Phase II duration = 30.63 (20.38 ; 34.86).
- Alternative scenario:
  - Power = P(Arm d_3 as winner in Phase III|Alternative scenario) = 0.812;
  - Phase II duration = 27.71 (14.08 ; 36.78).

Stroke trial

Null scenario

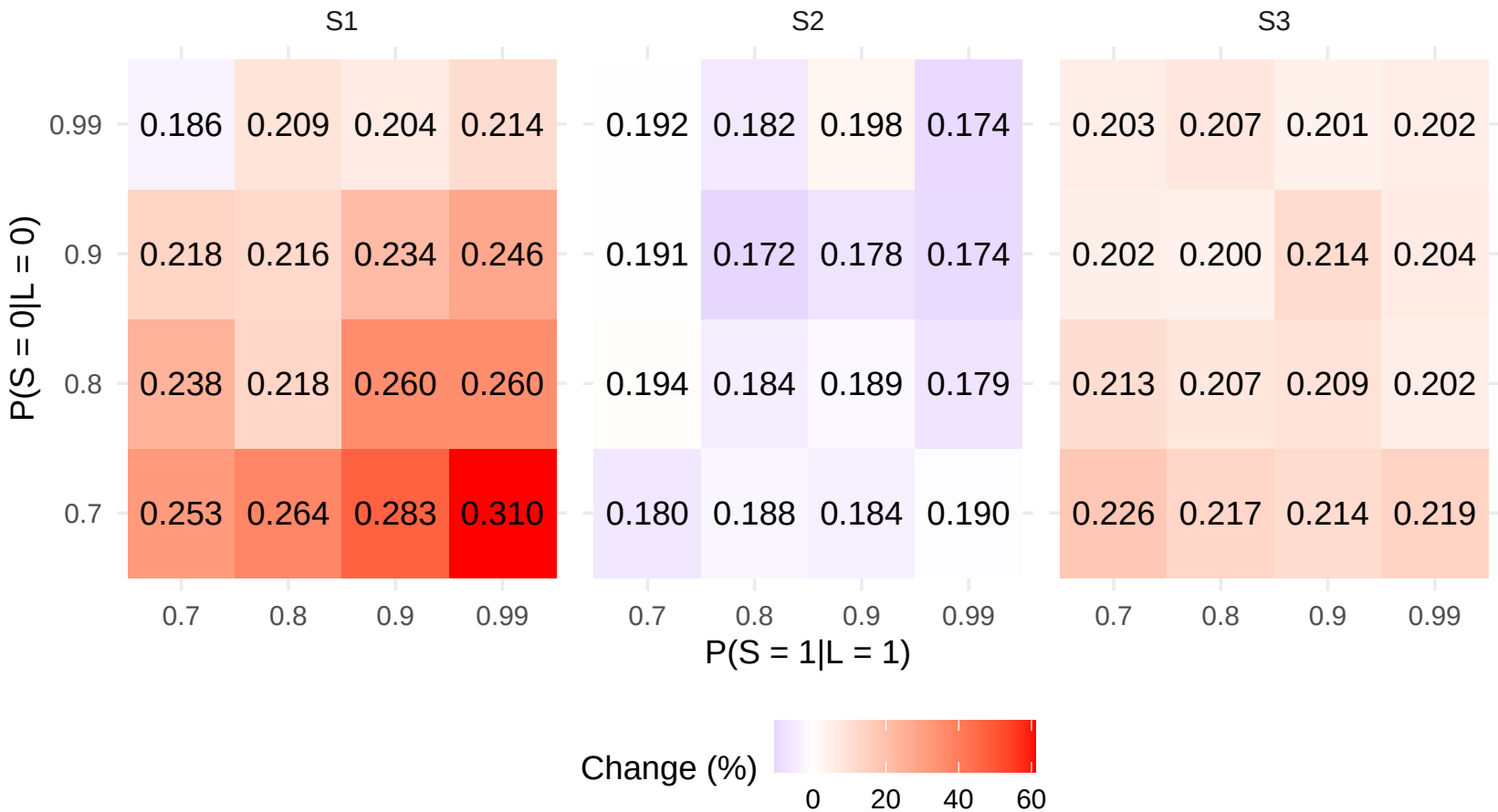


Figure: Transition probability

Stroke trial

Null scenario

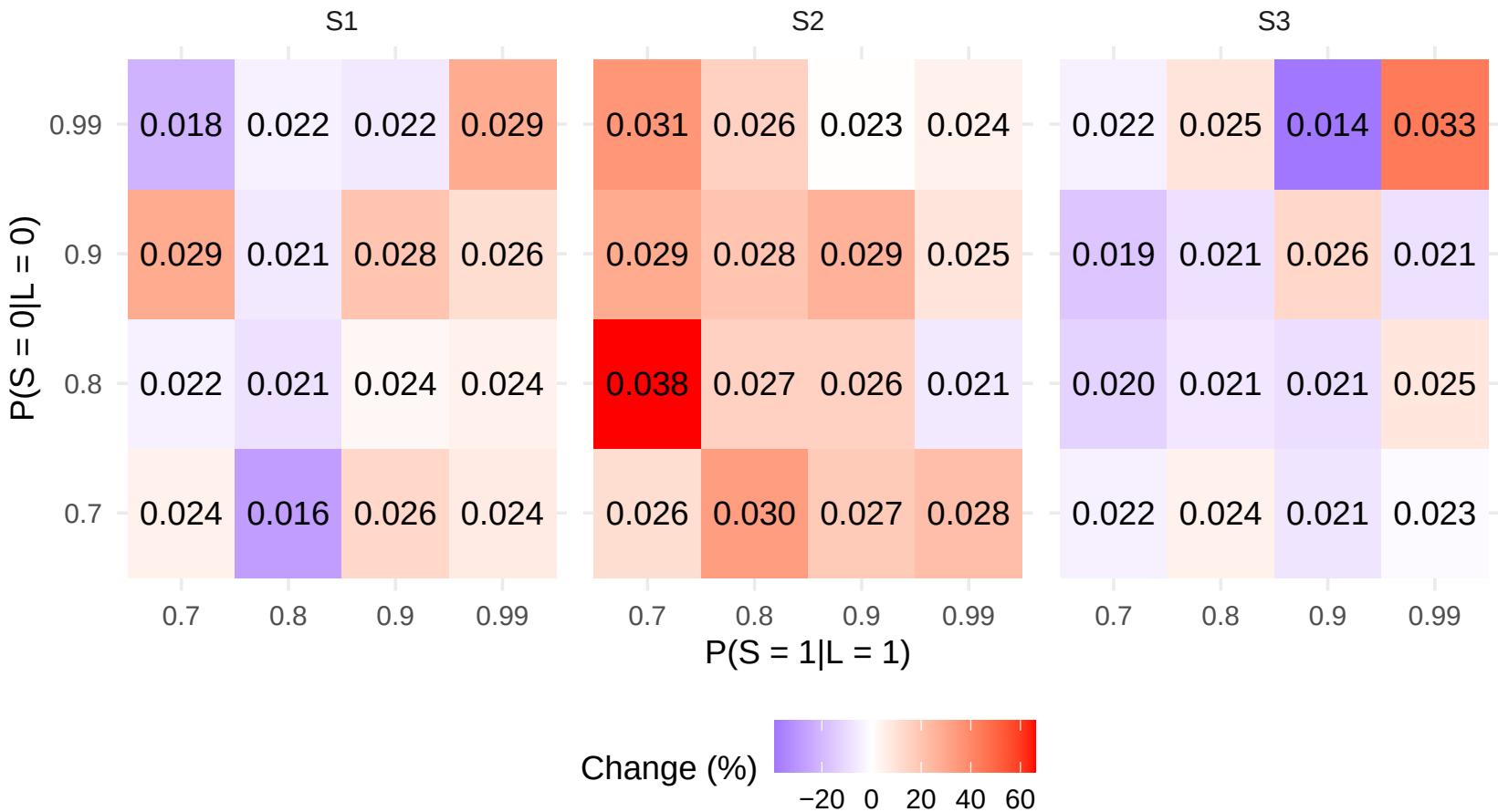


Figure: Type I Error

Stroke trial

Null scenario

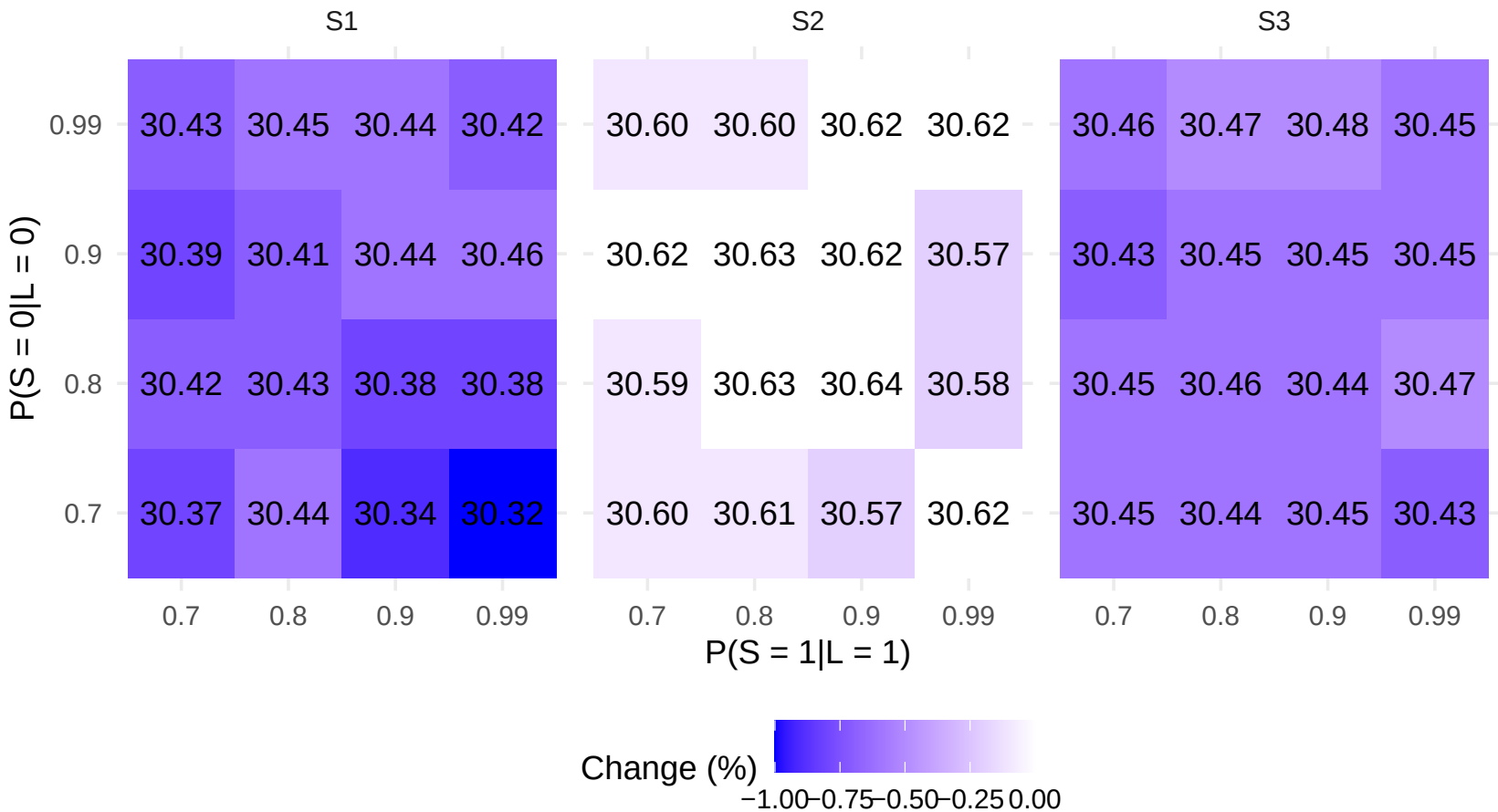


Figure: Phase II duration - Median

Stroke trial

Null scenario

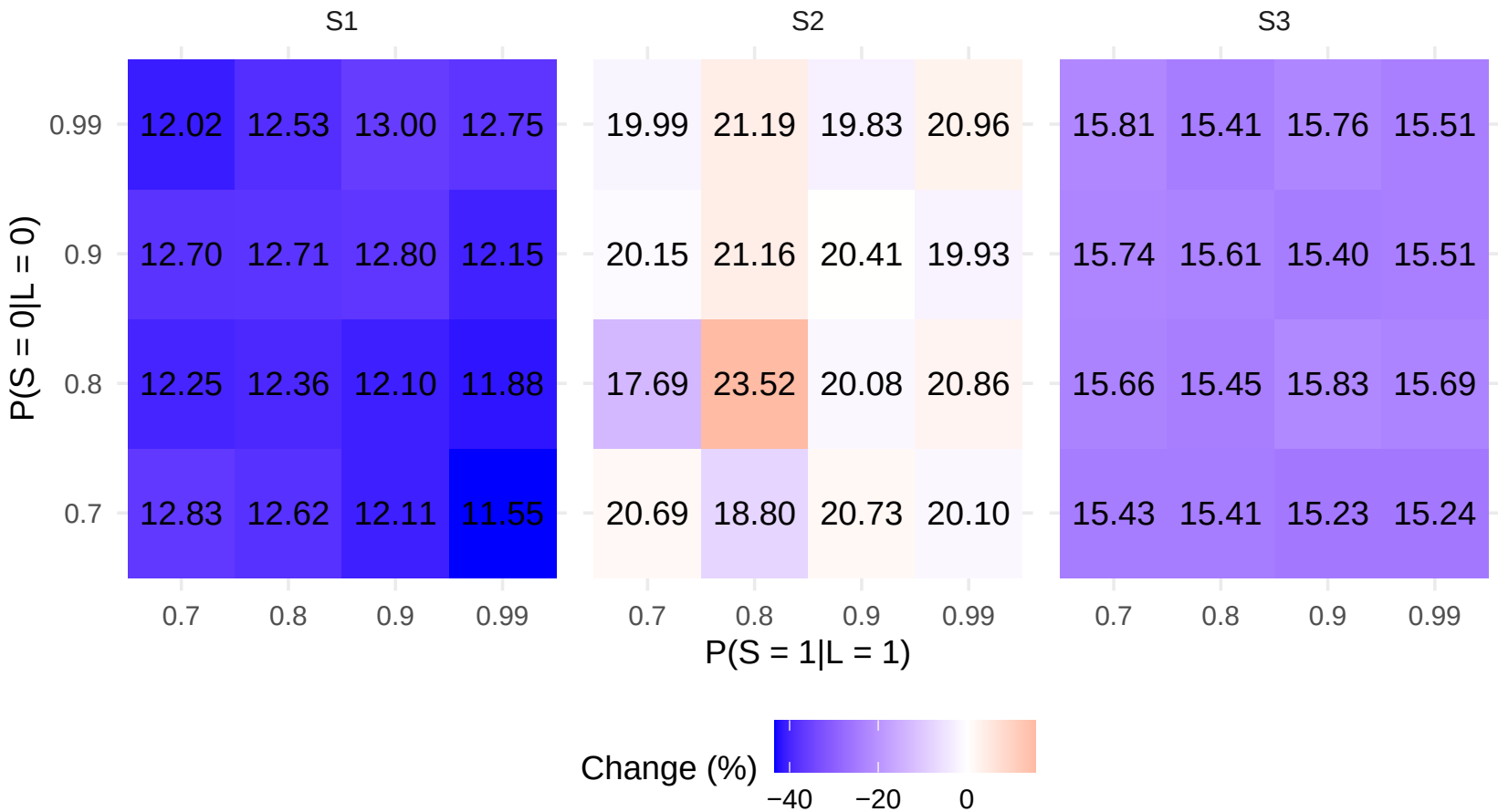


Figure: Phase II duration - Quantile 25%

Stroke trial

Null scenario

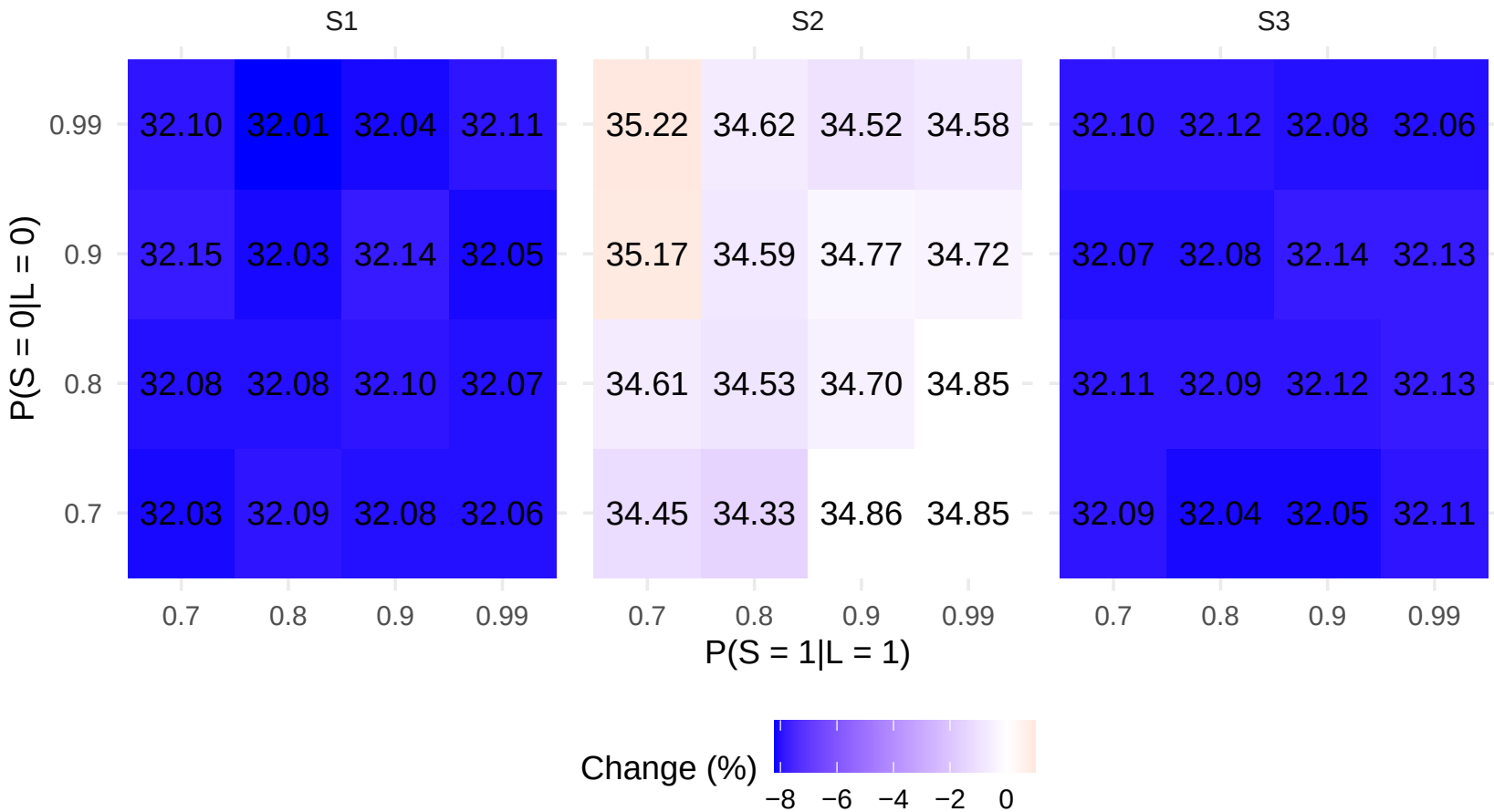


Figure: Phase II duration - Quantile 75%

Stroke trial

Alternative scenario

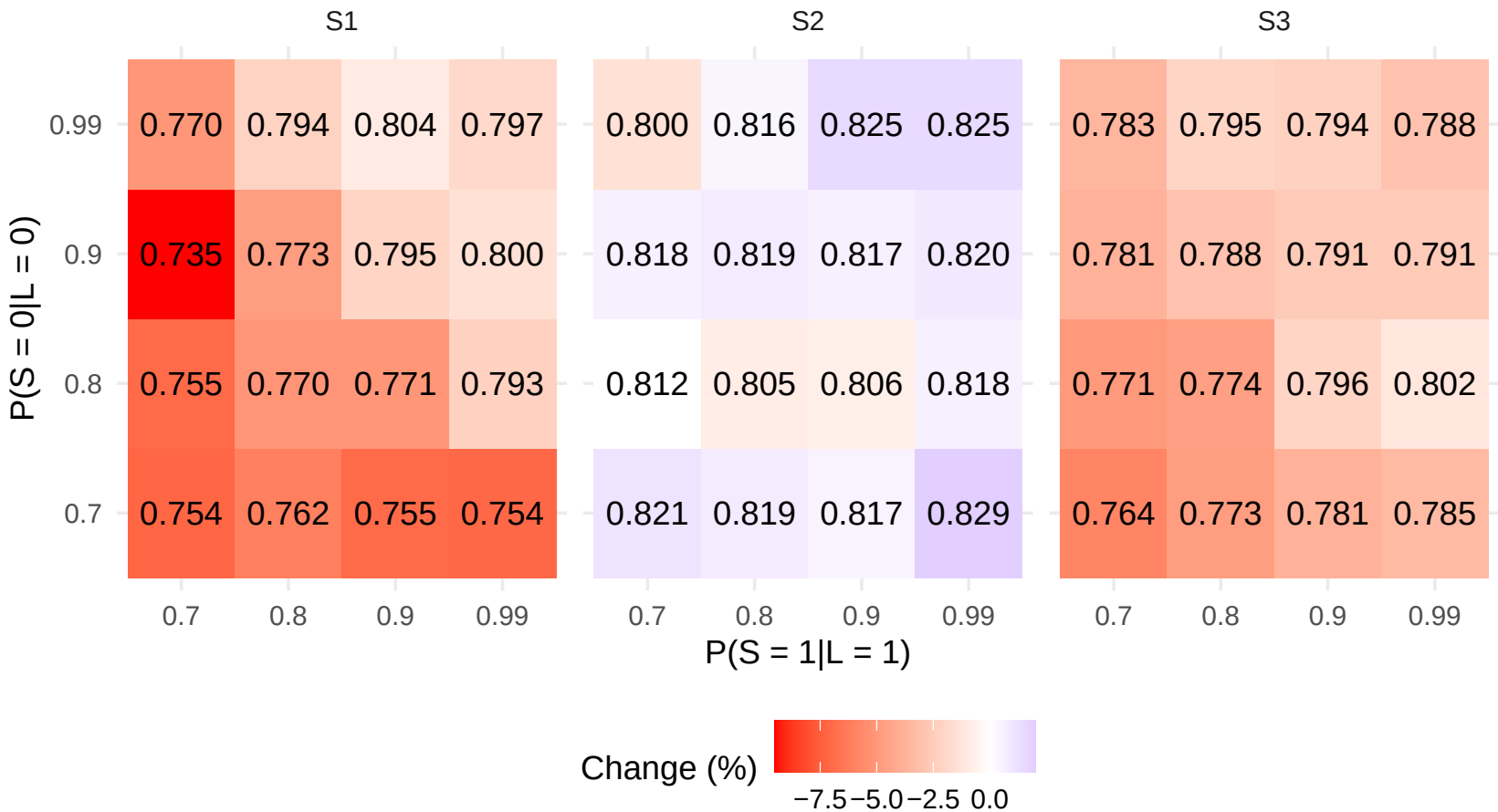


Figure: Power

Stroke trial

Alternative scenario

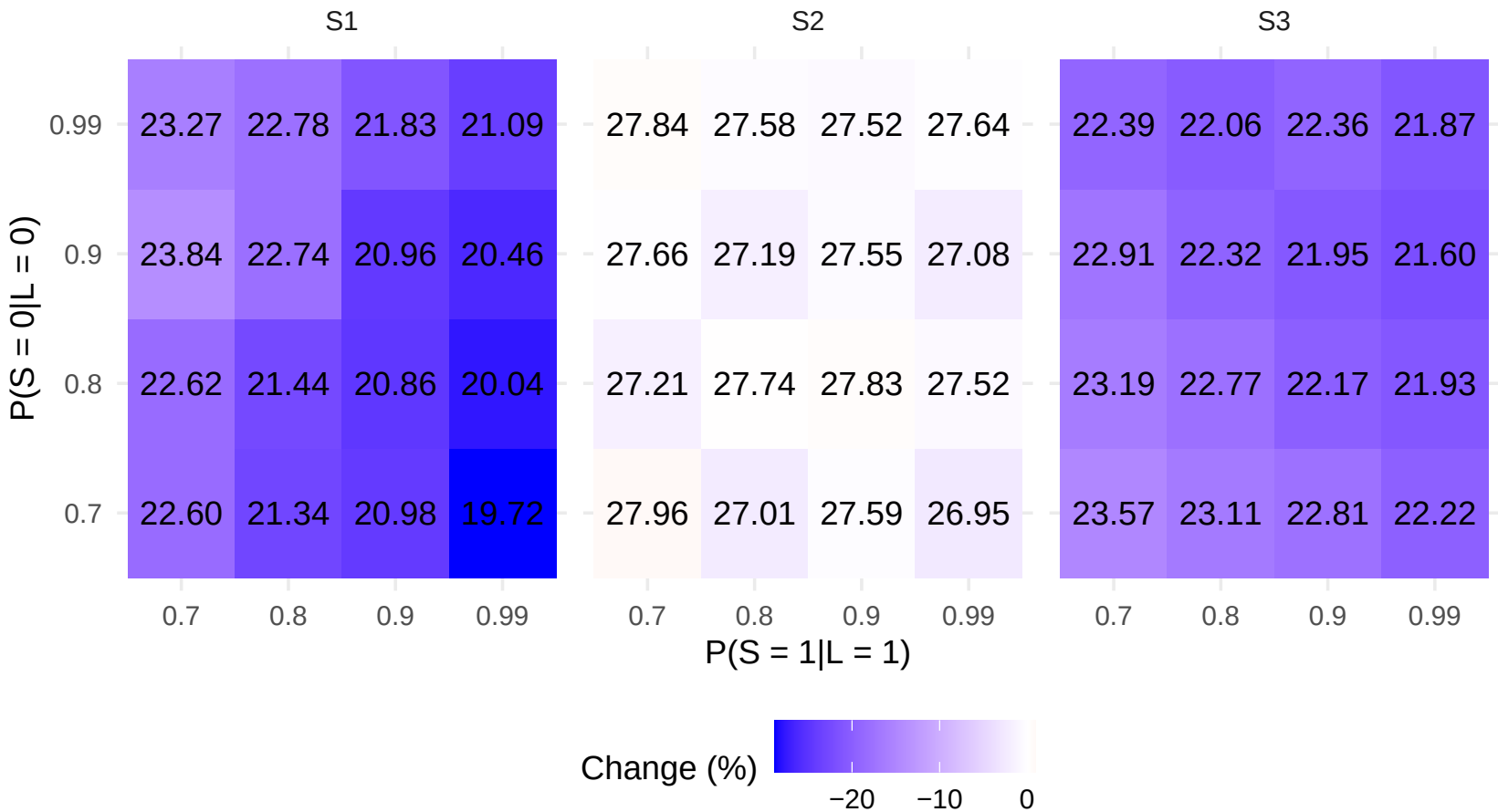


Figure: Phase II duration - Median

Stroke trial

Alternative scenario

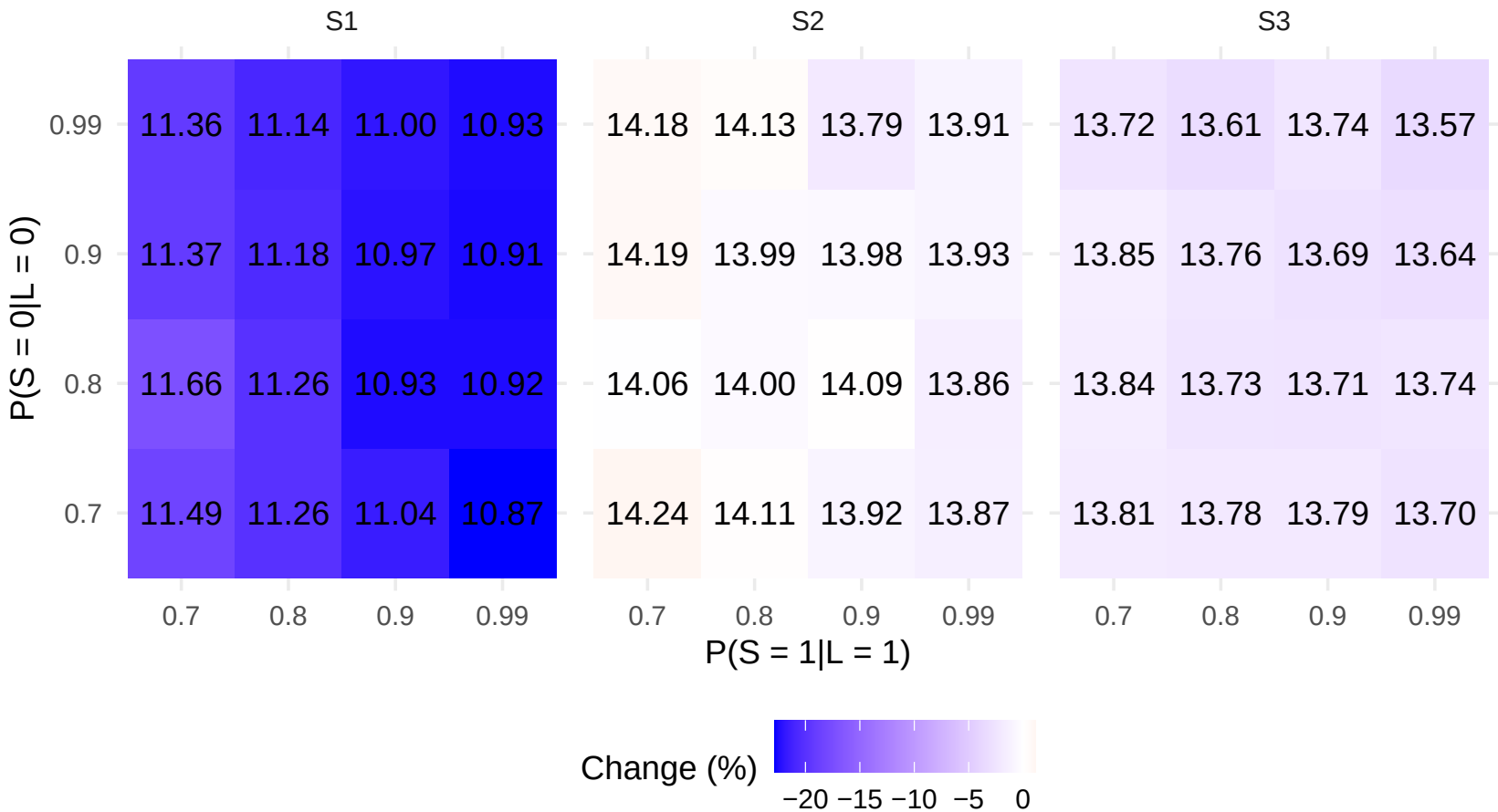


Figure: Phase II duration - Quantile 25%

Stroke trial

Alternative scenario

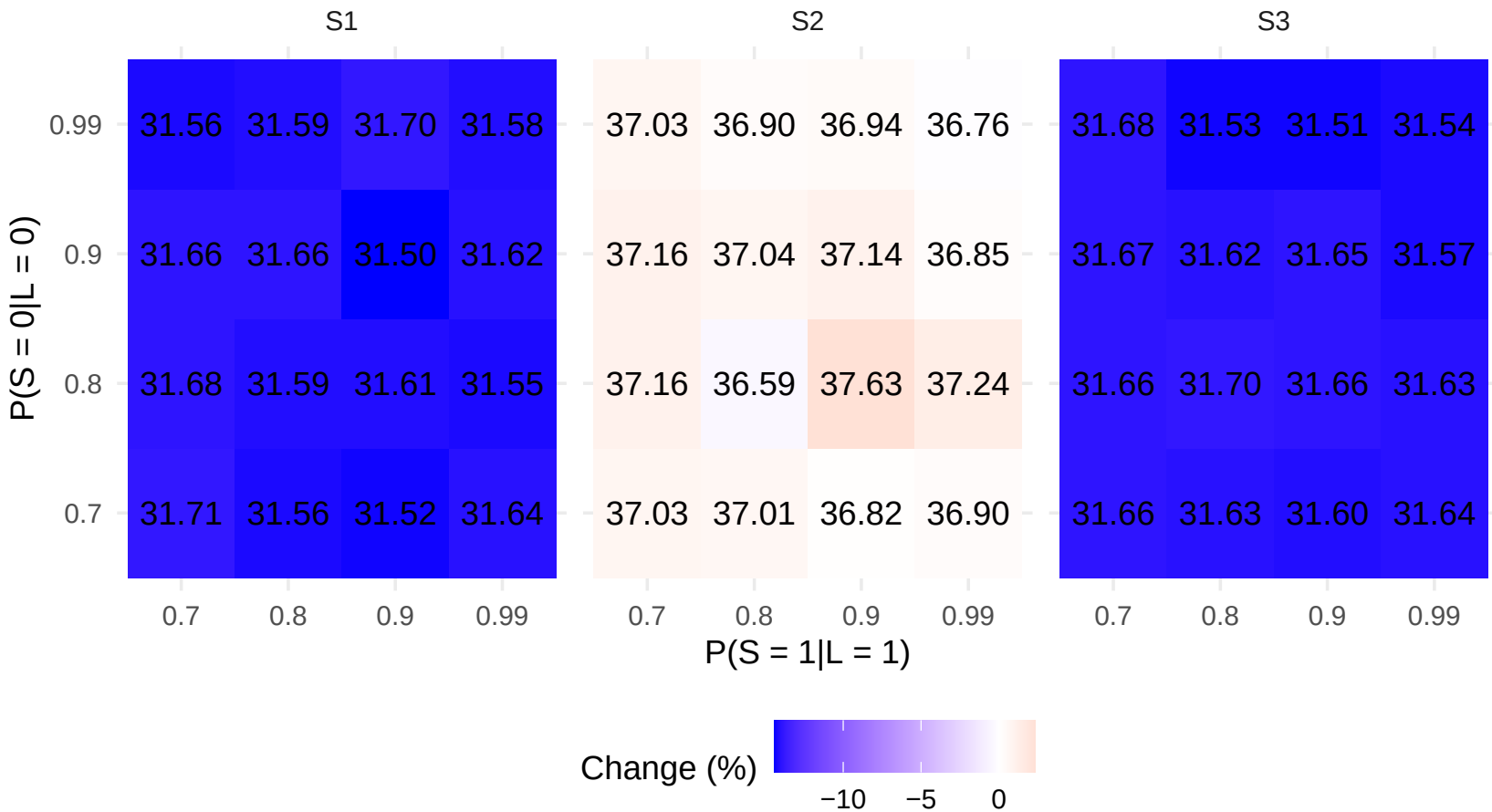


Figure: Phase II duration - Quantile 75%

Concluding Remarks

Strategies

- S_1 and S_3 shorten the phase II duration in around 4 months;
- S_1 inflates type I error up to 30%, while S_3 does not;
- S_1 and S_3 decreases power up to 8% as sensitivity and specificity decreases;
- S_2 slightly increases power and type I error, but does not decrease phase II duration.

Trial

- Modeling the misclassification should be done in case short-term endpoints are used as surrogates for long-term endpoints when designing a trial.
- Does the trade-off between shortening the trial in 4 months and increasing sample size to reach 80% power worth?

References I

Berry, D. A. and S. G. Eick (1995). Adaptive assignment versus balanced randomization in clinical trials: a decision analysis. *Statistics in medicine 14*(3), 231–246.

Bretz, F., H. Schmidli, F. König, A. Racine, and W. Maurer (2006). Confirmatory seamless phase ii/iii clinical trials with hypotheses selection at interim: general concepts. *Biometrical Journal: Journal of Mathematical Methods in Biosciences 48*(4), 623–634.

Daniel Paulino, C., P. Soares, and J. Neuhaus (2003). Binomial regression with misclassification. *Biometrics 59*(3), 670–675.

Eisele, J. R. (1994). The doubly adaptive biased coin design for sequential clinical trials. *Journal of Statistical Planning and Inference 38*(2), 249–261.

References II

Grayling, M. J., M. Dimairo, A. P. Mander, and T. F. Jaki (2019). A review of perspectives on the use of randomization in phase ii oncology trials. *JNCI: Journal of the National Cancer Institute 111*(12), 1255–1262.

Hey, S. P. and J. Kimmelman (2015). Are outcome-adaptive allocation trials ethical? *Clinical trials 12*(2), 102–106.

Huang, X., J. Ning, Y. Li, E. Estey, J.-P. Issa, and D. A. Berry (2009). Using short-term response information to facilitate adaptive randomization for survival clinical trials. *Statistics in medicine 28*(12), 1680–1689.

Inoue, L. Y., P. F. Thall, and D. A. Berry (2002). Seamlessly expanding a randomized phase ii trial to phase iii. *Biometrics 58*(4), 823–831.

Ivanova, A. (2003). A play-the-winner-type urn design with reduced variability. *Metrika 58*(1), 1–13.

References III

Jardim, D. L., E. S. Groves, P. P. Breitfeld, and R. Kurzrock (2017). Factors associated with failure of oncology drugs in late-stage clinical development: a systematic review. *Cancer treatment reviews 52*, 12–21.

Korn, E. L. and B. Freidlin (2011). Outcome-adaptive randomization: is it useful? *Journal of Clinical Oncology 29*(6), 771.

Maca, J., S. Bhattacharya, V. Dragalin, P. Gallo, and M. Krams (2006). Adaptive seamless phase ii/iii designs—background, operational aspects, and examples. *Drug Information Journal 40*(4), 463–473.

Meinzer, C., R. Martin, and J. I. Suarez (2017). Bayesian dose selection design for a binary outcome using restricted response adaptive randomization. *Trials 18*(1), 420.

References IV

Minnerup, J., H. Wersching, M. Schilling, and W. R. Schäbitz (2014). Analysis of early phase and subsequent phase iii stroke studies of neuroprotectants: outcomes and predictors for success. *Experimental & translational stroke medicine 6*(1), 2.

Nowacki, A. S., W. Zhao, and Y. Y. Palesch (2017). A surrogate-primary replacement algorithm for response-adaptive randomization in stroke clinical trials. *Statistical methods in medical research 26*(3), 1078–1092.

Stroke, T. A. I. R. I. (2001). Recommendations for clinical trial evaluation of acute stroke therapies. *Stroke 32*(7), 1598.

Tang, H., N. R. Foster, A. Grothey, S. M. Ansell, R. M. Goldberg, and D. J. Sargent (2010). Comparison of error rates in single-arm versus randomized phase ii cancer clinical trials. *Journal of Clinical Oncology 28*(11), 1936.

References V

Taylor, J. M., T. M. Braun, and Z. Li (2006). Comparing an experimental agent to a standard agent: relative merits of a one-arm or randomized two-arm phase ii design. *Clinical Trials 3*(4), 335–348.

Thall, P., P. Fox, and J. Wathen (2015). Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. *Annals of Oncology 26*(8), 1621–1628.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika 25*(3/4), 285–294.

Vickers, A. J., V. Ballen, and H. I. Scher (2007). Setting the bar in phase ii trials: the use of historical data for determining "go/no go" decision for definitive phase iii testing. *Clinical Cancer Research 13*(3), 972–976.

References VI

Wei, L. and S. Durham (1978). The randomized play-the-winner rule in medical trials. *Journal of the American Statistical Association 73*(364), 840–843.

Wilson, M. K., K. Karakasis, and A. M. Oza (2015). Outcomes and endpoints in trials of cancer treatment: the past, present, and future. *The Lancet Oncology 16*(1), e32–e42.

Wong, C. H., K. W. Siah, and A. W. Lo (2019). Estimation of clinical trial success rates and related parameters. *Biostatistics 20*(2), 273–286.