A Benchmark for Dose Finding Studies with Continuous Outcomes

Pavel Mozgunov, Thomas Jaki, Xavier Paoletti

Medical and Pharmaceutical Statistics Research Unit,
Department of Mathematics and Statistics, Lancaster University, UK

November 2, 2018

Acknowledgement: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567.

Motivation (Personal)

- A paper on Phase I/II design evaluating cumulative risk of toxicity over two treatment cycles & continuous activity. A typical result:

(Activity, Toxicity)	Dose 1	Dose 2	Dose 3	Dose 4
Scenario	(53,0.01)	(85,0.05)	(120,0.20)	(160,0.60)
Selection	0.0%	1.3%	**96.4%**	1.9%

Motivation (Personal)

- A paper on Phase I/II design evaluating cumulative risk of toxicity over two treatment cycles & continuous activity. A typical result:

(Activity, Toxicity)	Dose 1	Dose 2	Dose 3	Dose 4
Scenario	(53,0.01)	(85,0.05)	(120,0.20)	(160,0.60)
Selection	0.0%	1.3%	**96.4%**	1.9%

Impressive! But **no comparison** to alternative methods; reply: no alternatives.

Motivation (Personal)

- A paper on Phase I/II design evaluating cumulative risk of toxicity over two treatment cycles & continuous activity. A typical result:

(Activity, Toxicity)	Dose 1	Dose 2	Dose 3	Dose 4
Scenario	(53,0.01)	(85,0.05)	(120,0.20)	(160,0.60)
Selection	0.0%	1.3%	**96.4%**	1.9%

Impressive! But **no comparison** to alternative methods; reply: no alternatives.

- First to consider a continuous efficacy endpoint: Bekele and Shen (2005) Very high accuracy (over 80%) in all scenarios but, again, **no comparator**

Motivation (Personal)

- A paper on Phase I/II design evaluating cumulative risk of toxicity over two treatment cycles & continuous activity. A typical result:

(Activity, Toxicity)	Dose 1	Dose 2	Dose 3	Dose 4
Scenario	(53,0.01)	(85,0.05)	(120,0.20)	(160,0.60)
Selection	0.0%	1.3%	**96.4%**	1.9%

Impressive! But **no comparison** to alternative methods; reply: no alternatives.

- First to consider a continuous efficacy endpoint: Bekele and Shen (2005) Very high accuracy (over 80%) in all scenarios but, again, **no comparator**

"*We are not aware of any literature that formally incorporates a continuous activity outcome with toxicity ... in phase I/II clinical trials*"

Motivation (General)

The solution to the problem of **subjectively chosen scenarios:**
the *non-parametric optimal benchmark* by O'Quigley et al. (2002).

How "difficult" is to select the MTD in chosen scenario with binary responses?

Motivation (General)

The solution to the problem of **subjectively chosen scenarios:**
the *non-parametric optimal benchmark* by O'Quigley et al. (2002).

How "difficult" is to select the MTD in chosen scenario with binary responses?

Cheung (2014) generalized it for

- Phase I/II trials evaluating **binary** toxicity and efficacy endpoints
- Phase I trials with **multiple grades** of toxicities

Motivation (General)

The solution to the problem of **subjectively chosen scenarios:**
the *non-parametric optimal benchmark* by O'Quigley et al. (2002).

How "difficult" is to select the MTD in chosen scenario with binary responses?

Cheung (2014) generalized it for

- Phase I/II trials evaluating **binary** toxicity and efficacy endpoints
- Phase I trials with **multiple grades** of toxicities

There is a growing interest clinical trials involving **continuous endpoints**.

Motivation (General)

The solution to the problem of **subjectively chosen scenarios:**
the *non-parametric optimal benchmark* by O'Quigley et al. (2002).

How "difficult" is to select the MTD in chosen scenario with binary responses?

Cheung (2014) generalized it for

- Phase I/II trials evaluating **binary** toxicity and efficacy endpoints
- Phase I trials with **multiple grades** of toxicities

There is a growing interest clinical trials involving **continuous endpoints**.

Examples in oncology:

- Malignant glioma Phase I trial (Friedman et al., 1998)
- A cervical carcinoma Phase I/II trial (Hirakawa, 2012)

Examples outside of oncology:

- infectious diseases, respiratory diseases, cardiovascular diseases

Goal

To propose a simple benchmark, which can be applied to dose finding studies with continuous outcomes.

Goal

To propose a simple benchmark, which can be applied to dose finding studies with continuous outcomes.

Our proposal

- is a generalisation of the original benchmark;
- allows finding a benchmark for designs with multiple correlated outcomes and several treatment cycles;
- does not require any additional information other than already provided in the simulation study of a design.

Recall: The original benchmark

- Phase I, a **binary** toxicity outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} is a random variable: $y_{ij} = 1$ if patient i has experienced DLT at d_j
- $p_j = \mathbb{P}(Y_{ij} = 1)$ for $i = 1, \ldots, n$

Recall: The original benchmark

- Phase I, a **binary** toxicity outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} is a random variable: $y_{ij} = 1$ if patient i has experienced DLT at d_j
- $p_j = \mathbb{P}(Y_{ij} = 1)$ for $i = 1, \ldots, n$

The benchmark is based on the **complete information** concept, the vector of outcomes at **all dose levels** (in contrast to an actual trial).

The information about the DLT of patient i at each dose level is summarised in a single value $u_i \in (0, 1)$ **toxicity profile** (or **tolerance**), drawn from $\mathcal{U}(0, 1)$.

Recall: The original benchmark

- Phase I, a **binary** toxicity outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} is a random variable: $y_{ij} = 1$ if patient i has experienced DLT at d_j
- $p_j = \mathbb{P}(Y_{ij} = 1)$ for $i = 1, \ldots, n$

The benchmark is based on the **complete information** concept, the vector of outcomes at **all dose levels** (in contrast to an actual trial).

The information about the DLT of patient i at each dose level is summarised in a single value $u_i \in (0, 1)$ **toxicity profile** (or **tolerance**), drawn from $\mathcal{U}(0, 1)$.

For instance, $u_i = 0.3 \rightarrow$ patient i

- can tolerate doses d_j with $p_j \leq 0.3$;
- would experience a DLT if given dose $d_{j'}$ with $p_{j'} > 0.3$.

u_i is transformed to $y_{ij} = 0$ for doses with $p_j < 0.3$ and to $y_{ij} = 1$ otherwise.

Recall: The original benchmark

- Phase I, a **binary** toxicity outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} is a random variable: $y_{ij} = 1$ if patient i has experienced DLT at d_j
- $p_j = \mathbb{P}(Y_{ij} = 1)$ for $i = 1, \ldots, n$

The benchmark is based on the **complete information** concept, the vector of outcomes at **all dose levels** (in contrast to an actual trial).

The information about the DLT of patient i at each dose level is summarised in a single value $u_i \in (0, 1)$ **toxicity profile** (or **tolerance**), drawn from $\mathcal{U}(0, 1)$.

For instance, $u_i = 0.3 \rightarrow$ patient i

- can tolerate doses d_j with $p_j \leq 0.3$;
- would experience a DLT if given dose $d_{j'}$ with $p_{j'} > 0.3$.

u_i is transformed to $y_{ij} = 0$ for doses with $p_j < 0.3$ and to $y_{ij} = 1$ otherwise.

$$T(\mathbf{y}_j, \gamma) = \left| \frac{\sum_{i=1}^{n} y_{ij}}{n} - \gamma \right|.$$

Benchmark for Studies with Continuous Outcomes (I)

- Phase I with a **continuous** outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} at dose d_j for patient i with cumulative distr. function (CDF) $F_j(y)$.
- The goal is to find the target dose (TD) optimising criterion $T(\cdot)$.

Benchmark for Studies with Continuous Outcomes (I)

- Phase I with a **continuous** outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} at dose d_j for patient i with cumulative distr. function (CDF) $F_j(y)$.
- The goal is to find the target dose (TD) optimising criterion $T(\cdot)$.

Complete information: information about patient's profile is in $u_i \sim \mathcal{U}(0, 1)$.

Benchmark for Studies with Continuous Outcomes (I)

- Phase I with a **continuous** outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} at dose d_j for patient i with cumulative distr. function (CDF) $F_j(y)$.
- The goal is to find the target dose (TD) optimising criterion $T(\cdot)$.

Complete information: information about patient's profile is in $u_i \sim \mathcal{U}(0, 1)$.

How would patient with u_i respond to d_j with response's CDF F_j?

Benchmark for Studies with Continuous Outcomes (I)

- Phase I with a **continuous** outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} at dose d_j for patient i with cumulative distr. function (CDF) $F_j(y)$.
- The goal is to find the target dose (TD) optimising criterion $T(\cdot)$.

Complete information: information about patient's profile is in $u_i \sim \mathcal{U}(0,1)$.

How would patient with u_i respond to d_j with response's CDF F_j?

Probability integral transform

If $U \sim \mathcal{U}(0,1)$ is a uniform random variable on the unit interval, then F_j is the cumulative distribution function of a random variable $F_j^{-1}(U)$.

Benchmark for Studies with Continuous Outcomes (I)

- Phase I with a **continuous** outcome, n patients, doses $d_1, \ldots, d_m$.
- Y_{ij} at dose d_j for patient i with cumulative distr. function (CDF) $F_j(y)$.
- The goal is to find the target dose (TD) optimising criterion $T(\cdot)$.

Complete information: information about patient's profile is in $u_i \sim \mathcal{U}(0,1)$.

How would patient with u_i respond to d_j with response's CDF F_j?

Probability integral transform

If $U \sim \mathcal{U}(0,1)$ is a uniform random variable on the unit interval, then F_j is the cumulative distribution function of a random variable $F_j^{-1}(U)$.

For patient i with profile u_i, we apply the quantile transformation

$$y_{ij} = F_j^{-1}(u_i)$$

to obtain a continuous outcome that this patient would have at dose d_j.

Benchmark for Studies with Continuous Outcomes (II)

Different dose levels $d_1, \ldots, d_m$ are modelled by applying the quantile transformation using corresponding CDFs, $F_1, \ldots, F_m$.

Vector of responses $(y_{i1}, \ldots, y_{im})$, the **complete information** about patient i.

Benchmark for Studies with Continuous Outcomes (II)

Different dose levels $d_1, \ldots, d_m$ are modelled by applying the quantile transformation using corresponding CDFs, $F_1, \ldots, F_m$.

Vector of responses $(y_{i1}, \ldots, y_{im})$, the **complete information** about patient i.

The same procedure is repeated for all patients $i = 1, \ldots, n$, which, results in the vector of responses for each dose level

$$\mathbf{y}_j = (y_{1j}, \ldots, y_{nj})$$

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

Benchmark for Studies with Continuous Outcomes (II)

Different dose levels $d_1, \ldots, d_m$ are modelled by applying the quantile transformation using corresponding CDFs, $F_1, \ldots, F_m$.

Vector of responses $(y_{i1}, \ldots, y_{im})$, the **complete information** about patient i.

The same procedure is repeated for all patients $i = 1, \ldots, n$, which, results in the vector of responses for each dose level

$$\mathbf{y}_j = (y_{1j}, \ldots, y_{nj})$$

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

Q: How do *you* choose the CDFs $F_1, \ldots, F_m$?

Benchmark for Studies with Continuous Outcomes (II)

Different dose levels $d_1, \ldots, d_m$ are modelled by applying the quantile transformation using corresponding CDFs, $F_1, \ldots, F_m$.

Vector of responses $(y_{i1}, \ldots, y_{im})$, the **complete information** about patient i.

The same procedure is repeated for all patients $i = 1, \ldots, n$, which, results in the vector of responses for each dose level

$$\mathbf{y}_j = (y_{1j}, \ldots, y_{nj})$$

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

Q: How do *you* choose the CDFs $F_1, \ldots, F_m$?

A: I do not.

Phase I trial illustration (I)

The Phase I malignant glioma trial (Friedman et al., 1998) described by Wang and Ivanova (2015) measured a toxicity endpoint on a continuous scale.

Phase I trial illustration (I)

The Phase I malignant glioma trial (Friedman et al., 1998) described by Wang and Ivanova (2015) measured a toxicity endpoint on a continuous scale.

In one of the scenarios chosen by Wang and Ivanova (2015) to assess dose finding designs, it was assumed that a toxicity outcome Y_{ij} given dose level d_j has **normal distribution**

$$\mathcal{N}(0.1j, (0.1j)^2)$$

for $j = 1, \ldots, 6$.

Phase I trial illustration (I)

The Phase I malignant glioma trial (Friedman et al., 1998) described by Wang and Ivanova (2015) measured a toxicity endpoint on a continuous scale.

In one of the scenarios chosen by Wang and Ivanova (2015) to assess dose finding designs, it was assumed that a toxicity outcome Y_{ij} given dose level d_j has **normal distribution**

$$\mathcal{N}(0.1j, (0.1j)^2)$$

for $j = 1, \ldots, 6$.

Then, the CDF F_j is the CDF of a normal random variable with corresponding parameters

$$\Phi(\cdot,\ \mu_j\ =\ 0.1j\ ,\ \sigma^2\ =\ (0.1j)^2)$$

.

Phase I trial illustration (II)

Assume that the first patient has a toxicity profile $u_1 = 0.40$.

"How would patient 1 respond to d_1 corresponding to $\mathcal{N}(0.1, 0.01)$?"

Phase I trial illustration (II)

Assume that the first patient has a toxicity profile $u_1 = 0.40$.

"How would patient 1 respond to d_1 corresponding to $\mathcal{N}(0.1, 0.01)$?"

Applying the corresponding quantile transformation,

$$y_{11} = \Phi^{-1}(u_1 = 0.40, \mu_j = 0.1, \sigma^2 = 0.1^2) \approx 0.075$$

.

Subsequently, the complete information about patient 1 consists of the vector of responses at all dose levels $d_1, \ldots, d_6$,

$$(0.075, 0.149, 0.224, 0.299, 0.373, 0.448).$$

Phase I trial illustration (II)

Assume that the first patient has a toxicity profile $u_1 = 0.40$.

"How would patient 1 respond to d_1 corresponding to $\mathcal{N}(0.1, 0.01)$?"

Applying the corresponding quantile transformation,

$$y_{11} = \Phi^{-1}(u_1 = 0.40, \mu_j = 0.1, \sigma^2 = 0.1^2) \approx 0.075$$

.

Subsequently, the complete information about patient 1 consists of the vector of responses at all dose levels $d_1, \ldots, d_6$,

$$(0.075, 0.149, 0.224, 0.299, 0.373, 0.448).$$

Similarly, the complete information for patient with $u_2 = 0.25$

$$(0.033, 0.065, 0.098, 0.130, 0.163, 0.195)$$

and for patient with $u_3 = 0.92$

$$(0.241, 0.481, 0.722, 0.962, 1.203, 1.443).$$

Phase I trial illustration (III)

The objective of the design by Wang and Ivanova (2015): to choose the dose that maximises the probability of the average level of toxicity μ_j to be in the ε neighbourhood of γ

$$T(\mathbf{y}_j) = \int_{\gamma-\varepsilon}^{\gamma+\varepsilon} g_j(v|\mathbf{y}_j)\mathrm{d}v. \tag{1}$$

where $g_j(\cdot|\mathbf{y}_j)$ is the probability density function of μ_j given the data $\mathbf{y}_j$.

Using $\gamma = 0.1$, $\varepsilon = 0.01$, the complete information for tolerances u_1, u_2, u_3, and the density function of the Normal distribution:

$$T(\mathbf{y}_1) = 0.12;\ \ T(\mathbf{y}_2) = 0.04;\ \ T(\mathbf{y}_3) = 0.02;$$

$$T(\mathbf{y}_4) = 0.01;\ \ T(\mathbf{y}_5) = 0.01;\ T(\mathbf{y}_6) = 0.01$$

.

The procedure is repeated for $s = 1, \ldots, S$ simulated trials to obtain the proportion of correct selections.

Evaluation of the BDCO by Wang and Ivanova (2015)

Bayesian Design for Continuous Outcomes (BDCO)

- assumes that Y_{ij} at d_j for patient i has Normal distribution $\mathcal{N}(\mu_j, \sigma_j^2)$;
- μ_j is a random variable
- Based on the posterior of μ_j, BDCO is driven by

$$\pi_j = \mathbb{P}\left(\gamma - \varepsilon \leq \mu_j \leq \gamma + \varepsilon\right). \tag{2}$$

Scenarios:

- Six scenarios with six dose levels $d_1, \ldots, d_6$, $n = 36$
- (i) the case of equal variances, in which outcome Y_{ij} has normal distribution $\mathcal{N}(0.1j, 0.2^2)$
- (ii) the case of unequal variances corresponding to normal distributions $\mathcal{N}(0.1j, 0.1^2 j^2)$
- Different target values $\gamma = 0.1, 0.2, \ldots, 0.6$ were used

Evaluation of the BDCO. Results

Design	Variance	d_1	d_2	d_3	d_4	d_5	d_6
Scenario 1							
BDCO	Equal	**0.91**	0.10	0.00	0.00	0.00	0.00
Benchmark		**0.93**	0.07	0.00	0.00	0.00	0.00
Scenario 2							
BDCO	Equal	0.07	**0.86**	0.08	0.00	0.00	0.00
Benchmark		0.07	**0.87**	0.07	0.00	0.00	0.00
Scenario 3							
BDCO	Equal	0.00	0.07	**0.83**	0.09	0.00	0.00
Benchmark		0.00	0.07	**0.87**	0.07	0.00	0.00
Scenario 4							
BDCO	Equal	0.00	0.00	0.08	**0.81**	0.11	0.00
Benchmark		0.00	0.00	0.07	**0.87**	0.07	0.00
Scenario 5							
BDCO	Equal	0.00	0.00	0.00	0.09	**0.80**	0.11
Benchmark		0.00	0.00	0.00	0.07	**0.87**	0.07
Scenario 6							
BDCO	Equal	0.00	0.00	0.00	0.00	0.10	**0.90**
Benchmark		0.00	0.00	0.00	0.00	0.07	**0.93**

Evaluation of the BDCO. Results

Design	Variance	d_1	d_2	d_3	d_4	d_5	d_6
Scenario 1							
BDCO	Unequal	**0.97**	0.03	0.00	0.00	0.00	0.00
Benchmark		**0.96**	0.04	0.00	0.00	0.00	0.00
Scenario 2							
BDCO	Unequal	0.04	**0.84**	0.11	0.01	0.00	0.00
Benchmark		0.00	**0.88**	0.11	0.00	0.00	0.00
Scenario 3							
BDCO	Unequal	0.00	0.16	**0.65**	0.16	0.02	0.00
Benchmark		0.00	0.09	**0.74**	0.16	0.01	0.00
Scenario 4							
BDCO	Unequal	0.00	0.00	0.27	**0.50**	0.18	0.04
Benchmark		0.00	0.00	0.20	**0.59**	0.18	0.03
Scenario 5							
BDCO	Unequal	0.00	0.00	0.02	0.34	**0.45**	0.20
Benchmark		0.00	0.00	0.00	0.27	**0.48**	0.25
Scenario 6							
BDCO	Unequal	0.00	0.00	0.00	0.07	0.40	**0.54**
Benchmark		0.00	0.00	0.00	0.02	0.30	**0.67**

Benchmark for Multiple Endpoints

Consider a Phase I/II clinical trial with

- toxicity outcome $Y_{ij}^{(1)}$ having CDF $F_j^{(1)}$;
- efficacy outcome $Y_{ij}^{(2)}$ with CDFs $F_j^{(1)}$.

Benchmark for Multiple Endpoints

Consider a Phase I/II clinical trial with

- toxicity outcome $Y_{ij}^{(1)}$ having CDF $F_j^{(1)}$;
- efficacy outcome $Y_{ij}^{(2)}$ with CDFs $F_j^{(1)}$.

The **toxicity/efficacy profile** of patient i is given by $u_i^{(1)}, u_i^{(2)} \in (0, 1)$.

Benchmark for Multiple Endpoints

Consider a Phase I/II clinical trial with

- toxicity outcome $Y_{ij}^{(1)}$ having CDF $F_j^{(1)}$;
- efficacy outcome $Y_{ij}^{(2)}$ with CDFs $F_j^{(1)}$.

The **toxicity/efficacy profile** of patient i is given by $u_i^{(1)}, u_i^{(2)} \in (0, 1)$.

Generate a standard Normal vector $\left(x_i^{(1)}, x_i^{(2)}\right)$, $\mu = (0, 0)$ and $\Sigma = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$

Two correlated uniform RVs: $\left(u_i^{(1)}, u_i^{(2)}\right) = \left(\Phi(x_i^{(1)}), \Phi(x_i^{(2)})\right)$ (Tate, 1955).

Benchmark for Multiple Endpoints

Consider a Phase I/II clinical trial with

- toxicity outcome $Y_{ij}^{(1)}$ having CDF $F_j^{(1)}$;
- efficacy outcome $Y_{ij}^{(2)}$ with CDFs $F_j^{(1)}$.

The **toxicity/efficacy profile** of patient i is given by $u_i^{(1)}, u_i^{(2)} \in (0,1)$.

Generate a standard Normal vector $\left(x_i^{(1)}, x_i^{(2)}\right)$, $\mu = (0,0)$ and $\Sigma = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$

Two correlated uniform RVs: $\left(u_i^{(1)}, u_i^{(2)}\right) = \left(\Phi(x_i^{(1)}), \Phi(x_i^{(2)})\right)$ (Tate, 1955).

Quantile transformations are applied to $u_i^{(1)}$ and $u_i^{(2)}$ marginally. Values of the response for patient i at dose levels d_j are obtained as

$$y_{ij}^{(1)} = F_j^{-1^{(1)}}(u_i^{(1)}) \text{ and } y_{ij}^{(2)} = F_j^{-1^{(2)}}(u_i^{(2)})$$

.

Evaluation of the design by Bekele and Shen (2005)

The design:

- Bekele and Shen (2005) introduced a latent Normal random variable and a bivariate Normal distribution to model toxicity and efficacy jointly.
- Dose escalation/de-escalation decision rules are based on the posterior distribution of both toxicity and efficacy.

Scenarios:

- Total sample size was $N = 36$, $m = 4$ doses
- Efficacy outcome at d_j has Gamma distribution $\Gamma(\lambda_j\tau, \tau)$, $\tau = 0.1$
- DLT outcome has probability p_j.
- a weak association, $\rho = 0.25$
- Six scenarios specified by λ_j and p_j
- The target dose is the dose with the highest expected efficacy while being safe ($p_j < 0.35$) and efficacious ($\lambda_j > 5$).

Evaluation. Results

Scenario	Design	d_1	d_2	d_3	d_4	None
	(λ_j, p_j)	(25,0.01)	(70,0.10)	(115,0.25)	(127,0.60)	
Scenario 1	BS	0.00	0.03	**0.95**	0.02	
	Benchmark	0.00	0.09	**0.91**	0.00	0.00
	(λ_j, p_j)	(5,0.50)	(70,0.70)	(90,0.80)	(135,0.85)	
Scenario 2	BS	0.06	0.00	0.00	0.00	**0.94**
	Benchmark	0.02	0.00	0.00	0.00	**0.98**
	(λ_j, p_j)	(25,0.03)	(46,0.05)	(90,0.10)	(135,0.15)	
Scenario 3	BS	0.00	0.00	0.02	**0.98**	0.00
	Benchmark	0.00	0.00	0.01	**0.99**	0.00
	(λ_j, p_j)	(20,0.05)	(75,0.05)	(75,0.35)	(75,0.65)	
Scenario 4	BS	0.00	**0.83**	0.17	0.00	0.00
	Benchmark	0.00	**1.00**	0.00	0.00	0.00
	(λ_j, p_j)	(60,0.05)	(65,0.50)	(80,0.70)	(95,0.85)	
Scenario 5	BS	**0.94**	0.06	0.00	0.00	0.00
	Benchmark	**0.97**	0.03	0.00	0.00	0.00
	(λ_j, p_j)	(2,0.03)	(2,0.03)	(2,0.03)	(2,0.03)	
Scenario 6	BS	0.01	0.00	0.00	0.00	**0.99**
	Benchmark	0.01	0.00	0.00	0.00	**0.99**

Final Remarks

- The proposed benchmark is a **generalisation** of the original one.
- Continuous case open the door to **applications beyond dose-escalation** Phase I and Phase I/II clinical trials

Final Remarks

- The proposed benchmark is a **generalisation** of the original one.
- Continuous case open the door to **applications beyond dose-escalation** Phase I and Phase I/II clinical trials
- The benchmark is an evaluation tool and should be considered **together** with a design of interest

Final Remarks

- The proposed benchmark is a **generalisation** of the original one.
- Continuous case open the door to **applications beyond dose-escalation** Phase I and Phase I/II clinical trials
- The benchmark is an evaluation tool and should be considered **together** with a design of interest
- Nothing prevent us from using the generalised benchmark in **non-monotonic scenarios**
  Will it capture the setting of the unknown toxicity/efficacy ordering?

Further references

Paper: P. Mozgunov, T. Jaki and X. Paoletti. A benchmark for dose finding studies with continuous outcomes, *Biostatistics*, doi: 10.1093/biostatistics/kxy045

Code: Software in the form of R code is available on GitHub (`https://github.com/dose-finding/benchmark`).

References

Bekele, B. N. and Shen, Y. (2005) A bayesian approach to jointly modeling toxicity and biomarker expression in a phase i/ii dose-finding trial. *Biometrics*, **61**, 344–354.

Cheung, Y. K. (2014) Simple benchmark for complex dose finding studies. *Biometrics*, **70**, 389–397.

Friedman, H. S., Kokkinakis, D. M., Pluda, J., Friedman, A. H., Cokgor, I., Haglund, M. M., Ashley, D. M., Rich, J., Dolan, M. E., Pegg, A. E., Moschel, R. C., McLendon, R. E., Kerby, T., Herndon, J., Bigner, D. D. and Jr., S. C. S. (1998) Phase i trial of o6-benzylguanine for patients undergoing surgery for malignant glioma. *Journal of Clinical Oncology*, **16**, 3570–3575.

Hirakawa, A. (2012) An adaptive dose-finding approach for correlated bivariate binary and continuous outcomes in phase i oncology trials. *Statistics in Medicine*, **31**, 516–532.

O'Quigley, J., Paoletti, X. and Maccario, J. (2002) Non-parametric optimal design in dose finding studies. *Biostatistics*, **3**, 51–56.

Tate, R. F. (1955) The theory of correlation between two continuous variables when one is dichotomized. *Biometrika*, **42**, 205–216.

Wang, Y. and Ivanova, A. (2015) Dose finding with continuous outcome in phase i oncology trials. *Pharmaceutical Statistics*, **14**, 102–107.