Supporting Web Materials for Seamless phase I/II design for novel anticancer agents with competing disease progression

by Lucie Biard, Shing M. Lee, Bin Cheng

1 Calibration

Calibration was performed to obtain the dose skeleton for the simulation study using the indifference intervals approach (Lee and Cheung, 2009). We applied this approach to the toxicity exponential working model of the proposed design (CR-CRM). We implemented the algorithm proposed by Lee and Cheung (2009), estimating the probability of correct selection (PCS) with the CR-CRM design over a range of candidate widths of indifference intervals to define the dose skeleton. Eventually, we used the skeleton defined by the calibration interval width which yielded the best average probability of correctly selecting the best dose (PCS) across our simulation scenarios (set 1 and set 2 separately). As illustration, figure 1 reports the average PCS by interval half-width, across Set 1 scenarios and by sample size ($N = 50$ or 70), over 2000 simulations; figure S3 reports the PCS by interval half-width over 2000 simulations, for each scenario in Set 1.

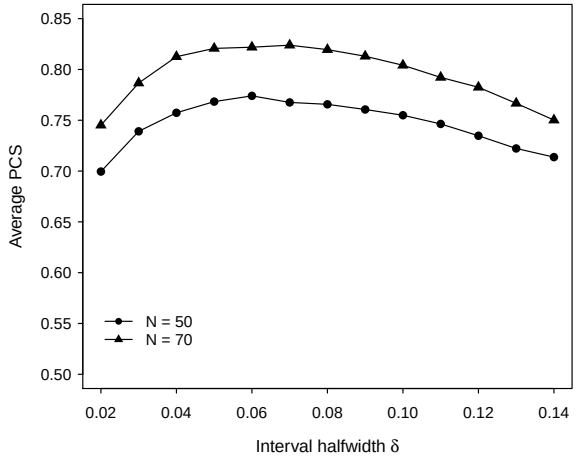

Figure S1: Average probability of correct selection (PCS) of the best dose across Set 1 scenarios, by calibration interval width, with $N = 25 + 25$ and $N = 35 + 35$ patients.

Figure S2 reports the average PCS by interval half-width across Set 1 scenarios depending on the sample size n_1 allocated to the first stage of the CR-CRM design, out of a total sample size of 50 patients, with $n_1 = 15, 20, 25$, corresponding to 30%, 40% and 50% of the total sample size (left panel); and similarly, with a total sample size of 70 patients (right panel): $n_1 = 20, 25, 30, 35$ corresponding respectively to 29%, 36%, 43% and 50% of the total sample size. An indifference interval half-width of 0.06 was consistently associated with a greater PCS, across our scenarios.

Figure S3 illustrates the variability in calibration profile across the different progression scenarios.

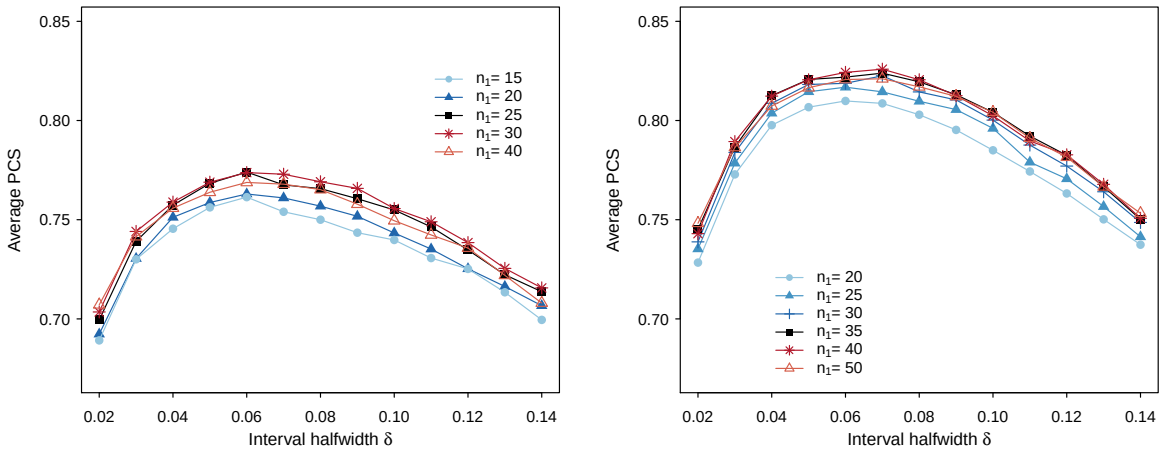

Figure S2: Average probability of correct selection (PCS) of best dose in scenarios in Set 1 with total sample sizes $N = 50$ (left panel) and $N = 70$ (right panel) with various sizes n_1 for stage 1.

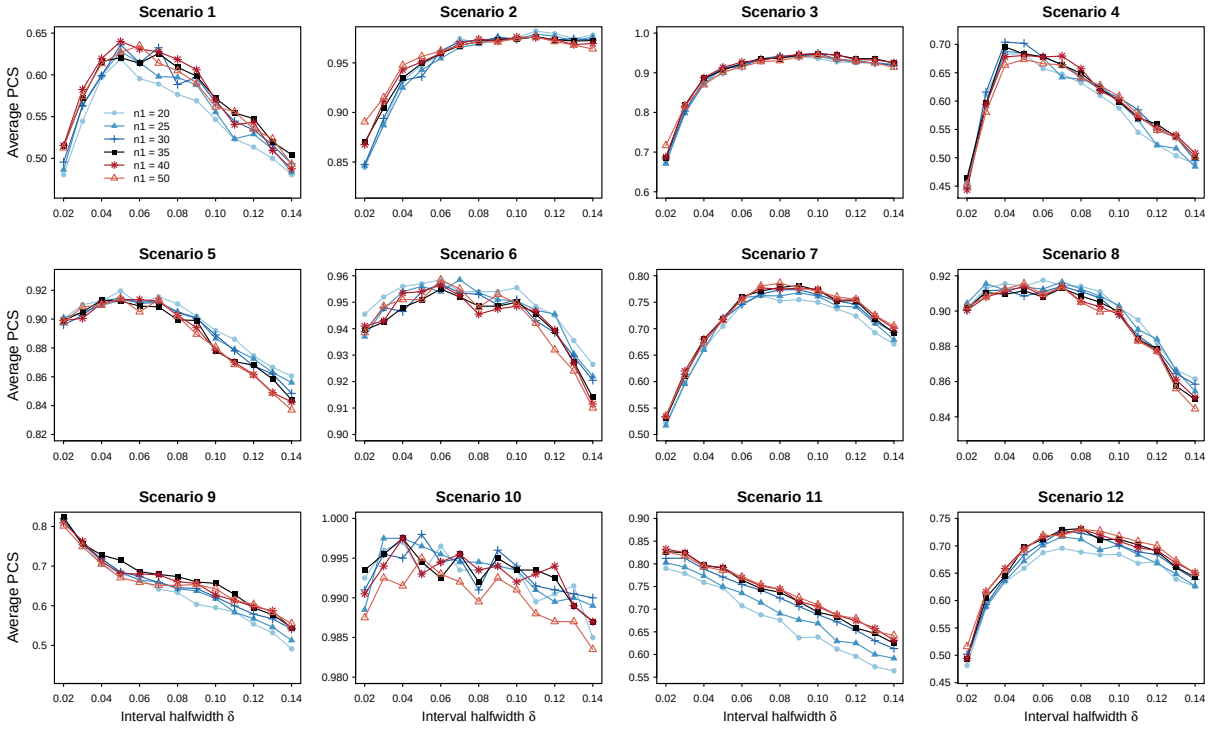

Figure S3: Average probability of correct selection (PCS) for best dose by interval width and by progression scenario, with $N = 70$ and various sizes n_1 for stage 1 (Set 1 scenarios: progression probability ranging from 0.60 to 0.27). Y-axis adapted to each scenario for comparison across n_1 values.

2 Additional simulation scenarios

In scenario Set 1 (Fig 1 of main manuscript), the effect size on the improvement in progression risk according to the dose level translates into a marginal cause specific incidence of progression at $t^\star$, $F_2(t^\star, d)$, ranging from 0.60 to 0.27. Given 5 candidate dose levels, possible values of $F_2(t^\star, d)$ were set equally-spaced within this range: 60.00, 51.75, 43.50, 35.25, 27.00%.

Figure S4 presents the second set, Set 2, of progression scenarios we evaluated. The effect size on the improvement in progression risk was smaller: from $F_2(t^\star, d) = 0.60$ to 0.37. In Set 2, the possible values of $F_2(t^\star, d)$ were also set equally-spaced within this range: 60.00, 54.25, 48.50, 42.75, 37.00%.

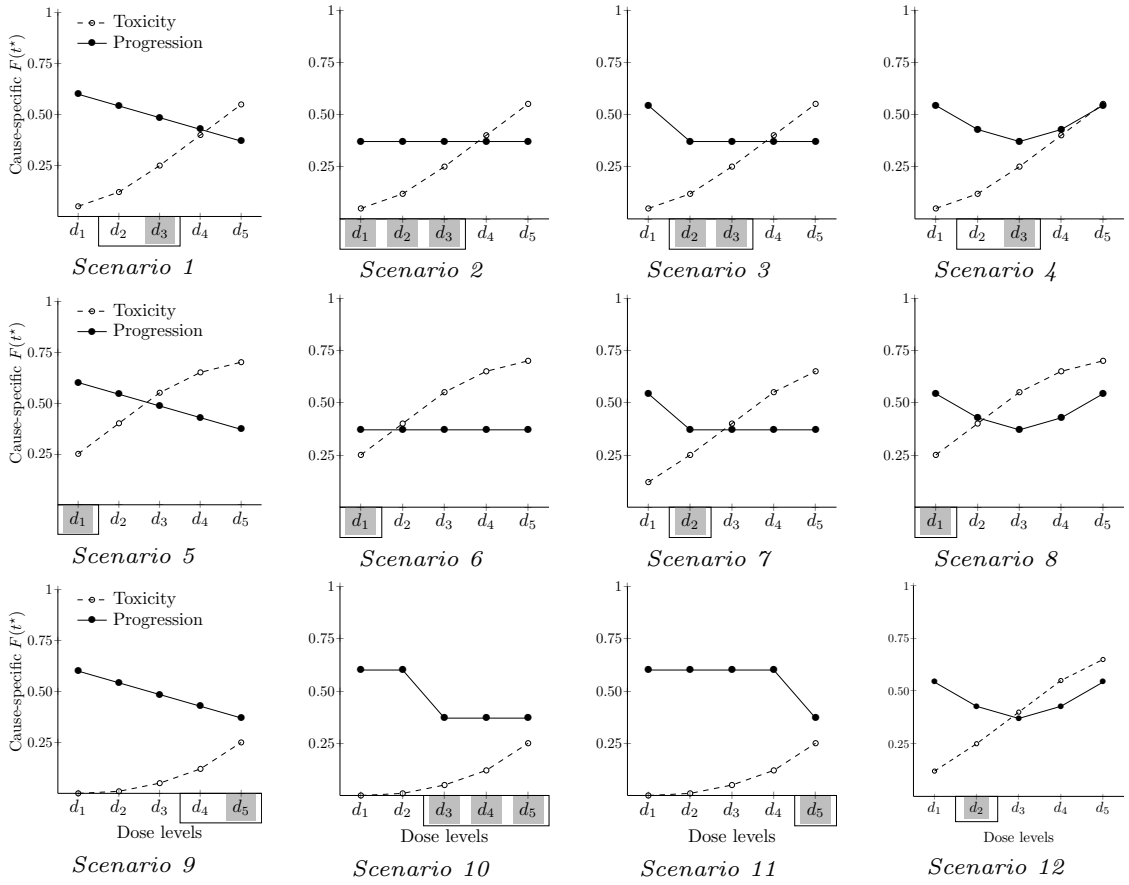

Figure S4: Simulation scenarios: Set 2, with progression cause-specific incidence ranging from 0.60 to 0.37 across doses. For each scenario, the best dose level(s) are shaded in gray and acceptable doses level are squared in black.

3 Data generation

3.1 Simulation of time-to-event data with competing toxicity and progression

Time-to-event data with competing toxicity and progression in the setting of a dose-finding clinical trial was generated following Beyersmann et al. (2009)'s approach, with Weibull distributions for time-varying cause-specific hazards.

We summarize the simulation procedure below. R software notations for the Weibull distribution are used: let a_1 be the shape parameter and b_1 the scale parameter for the toxicity cause-specific hazard $h_1(t)$, let a_2 be the shape parameter and b_2 the scale parameter for the progression cause-specific hazard $h_2(t)$. To generate a patient's outcome under a dose level d, the steps are as follows (Beyersmann et al., 2009):

1. Define $h_1(t) = \frac{a_1}{b_1}\left(\frac{t}{b_1}\right)^{a_T-1}$ and $h_2(t) = \frac{a_2}{b_2}\left(\frac{t}{b_2}\right)^{a_2-1}$.
   The cumulative hazards are $H_1(t) = \left(\frac{t}{b_1}\right)^{a_1}$ and $H_2(t) = \left(\frac{t}{b_2}\right)^{a_2}$.
2. Generate all-cause event times T with all-cause hazard function $h_{all}(t) = h_1(t) + h_2(t)$ (see details below)
3. Given time T, obtain the observed type of event using a Bernoulli distribution with probability $\frac{h_1(T)}{h_1(T)+h_2(T)}$ for toxicity
4. Apply administrative censoring if $T > t^\star$, to obtain observations corresponding to clinical trial data with pre-specified observation window $t^\star$. We assume no additional censoring process in the specific setting of early phase trials in oncology.

For step (b), we use the event-free survival function, defined as:

$$S(t) = \exp\left\{-\sum_{k=1}^{2} H_k(t)\right\} = \exp\left[-\left(\frac{t}{b_1}\right)^{a_1} - \left(\frac{t}{b_2}\right)^{a_2}\right]$$

.

For each patient i, we generate the all-cause event times T_i by inverse transform sampling from $S(t)$: we draw an observation u_i from a uniform distribution on $[0;1]$, and then we obtain $T_i = S^{-1}(u_i)$, solving the following equation:

$$\exp\left[-\left(\frac{T_i}{b_1}\right)^{a_1} - \left(\frac{T_i}{b_2}\right)^{a_2}\right] = u_i.$$

The main simulations were performed with constant hazard for toxicity and progression, setting Weibull shape parameters a_1 and a_2 equal to 1. At each dose level, given the shape parameters, a_1 and a_2, scale parameters, b_1 and b_2, were obtained respectively by solving the equations for the cause-specific cumulative incidences, $F_1(t^\star, d)$ and $F_2(t^\star, d)$, specified according to the scenarios (see figure S1): $F_k(t^\star, d) = 1 - \exp\{-H_k(t^\star, d)\}$, $k = 1, 2$.

As illustrative example, scale parameters, b_1 and b_2, for scenario 1 are displayed in table 1: they were obtained given shapes values, a_1 and a_2 respectively, and according to pre-specified cause-specific cumulative incidences of events at $t^\star = 8$ (by dose level). For sensitivity analyses, we generated observations with time-varying cause-specific hazards to simulate late-onset or early-onset toxicity or progression: we used a shape parameter of 0.3 to obtain time-decreasing hazard, and a shape=3 to obtain time-increasing hazard (table 1). Figure S5 illustrates the obtained different hazards, given a cause-specific cumulative incidence of 0.25 at $t^\star = 8$ weeks.

Table S 1: Time-to-event simulations with competing toxicity and progression: example of Weibull parameters by dose level in Scenario 1, with constant (shape=1), decreasing (shape=0.3) or increasing hazard (shape=3), and $t^\star = 8$.

	Toxicity			Progression		
Dose level	$F_1(t^\star, d)$	Shape a_1	Scale b_1	$F_2(t^\star, d)$	Shape a_2	Scale b_2
1	5.00%	1	156.0	60.00%	1	8.7
		0.3	159547.8		0.3	10.7
		3	21.5		3	3.2
2	12.00%	1	62.6	51.75%	1	11.0
		0.3	7602.3		0.3	23.0
		3	15.9		3	8.9
3	25.00%	1	27.8	43.50%	1	14.0
		0.3	509.0		0.3	51.8
		3	12.1		3	9.6
4	40.00%	1	15.7	35.25%	1	18.4
		0.3	75.1		0.3	128.6
		3	10.0		3	10.6
5	55.00%	1	10.0	27.00%	1	25.4
		0.3	16.9		0.3	377.3
		3	8.6		3	11.8

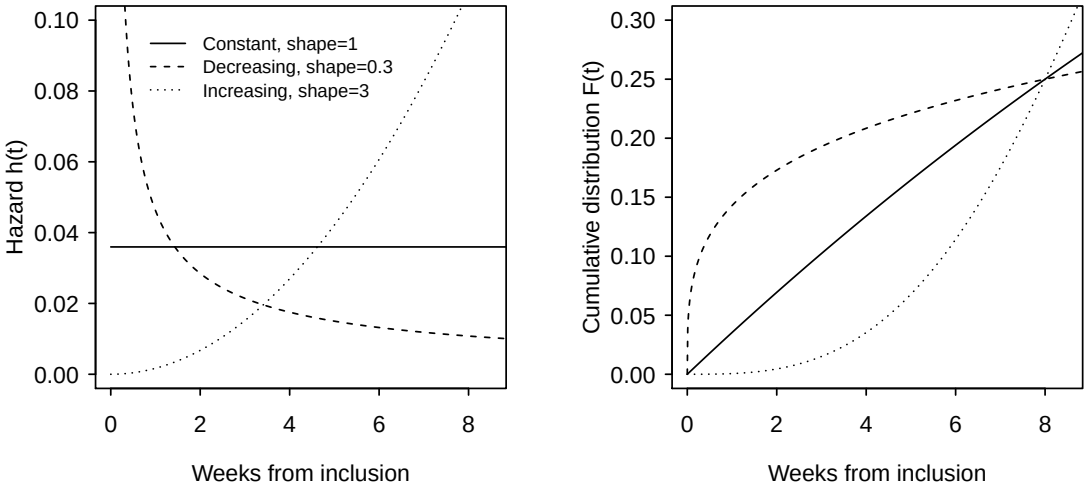

Figure S5: Simulated Weibull hazard functions $h(t)$ (left panel) and corresponding cumulative distribution $F(t)$ (right panel) over the observation window $t^\star = 8$: constant (shape=1) or time-varying (decreasing (shape=0.3), increasing (shape=3), with pre-specified $F(t^\star) = 0.25$.

3.2 Simulation of correlated time to toxicity and time to progression

For sensitivity analyses, we generated trial data under Clayton's model, for correlated time to toxicity and time to progression, as proposed in Yuan and Yin (2009).

This approach consists in generating correlated pairs of time to toxicity and time to progression, for each patient. Denote t_T the time to toxicity and t_P the time to progression. Clayton's model defines the joint density for (t_T, t_P) as:

$$f(t_T, t_P) = \frac{\phi+1}{\phi}\left\{S_T(t_T)^{-1/\phi} + S_P(t_P)^{-1/\phi} - 1\right\}^{-\phi-2} f_T(t_T) f_P(t_P) \left\{S_T(t_T) S_P(t_P)\right\}^{-1/\phi-1},$$

where $f_T(t_T)$ and $f_P(t_P)$ are the marginal density functions for t_T and t_P respectively, and ϕ represents the correlation between time to toxicity and time to progression. The progression-free survival conditional on the time to toxicity can be written as follows:

$$S_{P|T}(t_P|t_T) = \{S_T(t_T)^{-1/\phi} + S_P(t_P)^{-1/\phi} - 1\}^{-\phi-1} S_T(t_T)^{-(\phi+1)/\phi}.$$

Given $a = S_T(t_T)^{-1/\phi} - 1$ and $b = S_T(t_T)^{-(\phi+1)/\phi}$, we obtain:

$$\begin{aligned}
S_{P|T}(t_P|t_T) &= \{a + S_P(t_P)^{-1/\phi}\}^{-\phi-1} \times b \\
\frac{S_{P|T}(t_P|t_T)}{b} &= \{a + S_P(t_P)^{-1/\phi}\}^{-\phi-1} \\
\left\{\frac{S_{P|T}(t_P|t_T)}{b}\right\}^{1/(-\phi-1)} &= a + S_P(t_P)^{-1/\phi} \\
\left\{\frac{S_{P|T}(t_P|t_T)}{b}\right\}^{1/(-\phi-1)} - a &= S_P(t_P)^{-1/\phi} \\
S_P(t_P) &= \left[\left\{\frac{S_{P|T}(t_P|t_T)}{b}\right\}^{1/(-\phi-1)} - a\right]^{-\phi}
\end{aligned}$$

We obtain the bivariate (t_T, t_P) by inverse transform sampling: we draw a two independent random observations (u_1, u_2) from the uniform distribution $(0, 1)$ and generate the marginal time to toxicity as $t_T = S_T^{-1}(u_1)$ and t_P as $t_P = S_{P|T}^{-1}(u_2)$ (Yuan and Yin, 2009).

In particular, we assumed exponential marginal survival functions, $S_T(t_T)$ and $S_P(t_P)$, with constant hazards. These hazards were defined by dose level, according to the desired scenarios in terms of marginal incidence at time $t^\star$, similarly to the main simulation study (see, for instance, section 3.1 above and table 1). For each patient, we considered only the first occurring event within the observation window, if any, could be observed in our setting of competing events (both types of event resulting in trial discontinuation).

4 Benchmark method

To obtain benchmark performances for our trial objective given a trial sample size N and an observation window $t^\star$, we applied the non parametric benchmark method for dose-finding trials (O'Quigley et al., 2002). In the setting of time-to-event competing endpoints with a bivariate dose-finding objective relying on the cause-specific risks of DLT and progression, we consider an approach for complex designs (Cheung, 2014).

We generated the complete outcome profile of trial patients using the simulation approach for competing risks we described above (Beyersmann et al., 2009). While only the time-to-event outcome under the assigned dose was used as information in the CR-CRM design, the benchmark made use of the complete profiles of outcomes under all dose levels.

To warrant consistency of outcomes across dose levels, we first defined a tolerance profile for each patient (O'Quigley et al., 2002; Cheung, 2014). In the setting of time-to-event data with competing risks, all information (time-to-event and type of event) for a given patient could be summarized in the

profile $(u_t, u_K) \sim {}^{\text{iid}}\text{Uniform}(0,1)$, where u_t models the tolerance related to the time to all-cause event T_i, and u_K the tolerance to the type of event.

For each patient i, given cause-specific hazards for toxicity and progression defined after the simulation scenarios, we generated the time to any event $T_i(d)$, for each dose level, as a continuous variable, using inverse transform sampling (Beyersmann et al., 2009; Mozgunov et al., 2020): $T_i(d) = S^{-1}(u_t, d)$ where $S(., d)$ is the event-free survival function at dose d. At each dose level, the type of event was then obtained using the profile value u_K and inverse transform sampling on the Bernoulli distribution with probability $\frac{h_2(t^\star)}{h_1(t^\star)+h_2(t^\star)}$ for progression ($K = 2$; and otherwise toxicity, $K = 1$). Last, administrative censoring at time $t^\star$ was applied to mimick trial data.

We computed non parametric estimates of the marginal cause-specific incidences by dose level, $\hat{F}_1(t^\star, d), \hat{F}_2(t^\star, d)$, under the working assumption of independence between toxicity and progression, applying the Kaplan Meier estimator to the generated complete information. Benchmark performances estimates of selecting each dose were then computed based on the trial objective, plugging $\hat{F}_1(t^\star, d), \hat{F}_2(t^\star, d)$ in the dose-finding objectives for best and good doses (d_ν and D_ν).

5 Sensitivity simulations

Figure S6 reports the estimated probability of correctly selecting the best dose depending on the total sample size, $N = 50$ or 70, with $n_1 = 0.5N$, with the proposed CR-CRM design and corresponding benchmark and Wages and Tait's design (W&T) performances.

Table S2 reports the results of the proposed design using a modified definition of the tolerable dose set (modified CR-CRM, mCR-CRM), including dose levels with a DLT probability by $t^\star$ lower than or equal to the target probability (25%).

Figure S7 reports the estimated probability of correctly selecting the best dose depending on the effect size on the progression risk, comparing Set 1 and Set 2 of progression scenarios, with $N = 70$ and $n_1 = 0.5N$, with the proposed CR-CRM design and corresponding benchmark and Wages and Tait's design (W&T) performances.

Figures S8 and S9 report the estimated probability of correctly selecting the best dose and of overdose selection depending on the type of true hazards: constant, decreasing, increasing or correlated (Clayton's model), with $N = 70$ and $n_1 = 0.5N$.

Table S2: Simulation results with the modified CR-CRM algorithm (mCR-CRM), with modified definition of the tolerable set: percent of selecting good, best or toxic dose. Simulated trials with targeted maximum toxicity = 25% and $N = 35 + 35$. (n/a : not applicable)

				Probability of picking (percent)		
Scenario	MTD	Good	Best	Good dose	Best dose	Toxic dose
S1	3	2,3	3	87	47	2
S2	3	1,2,3	1,2,3	99	99	1
S3	3	2,3	2,3	96	96	1
S4	3	2,3	3	95	46	1
S5	1	1	1	97	97	1
S6	1	1	1	99	99	1
S7	2	2	2	52	52	1
S8	1	1	1	97	97	3
S9	5	4,5	5	92	49	n/a
S10	5	3,4,5	3,4,5	99	99	n/a
S11	5	5	5	52	52	n/a
S12	2	2	2	50	50	1

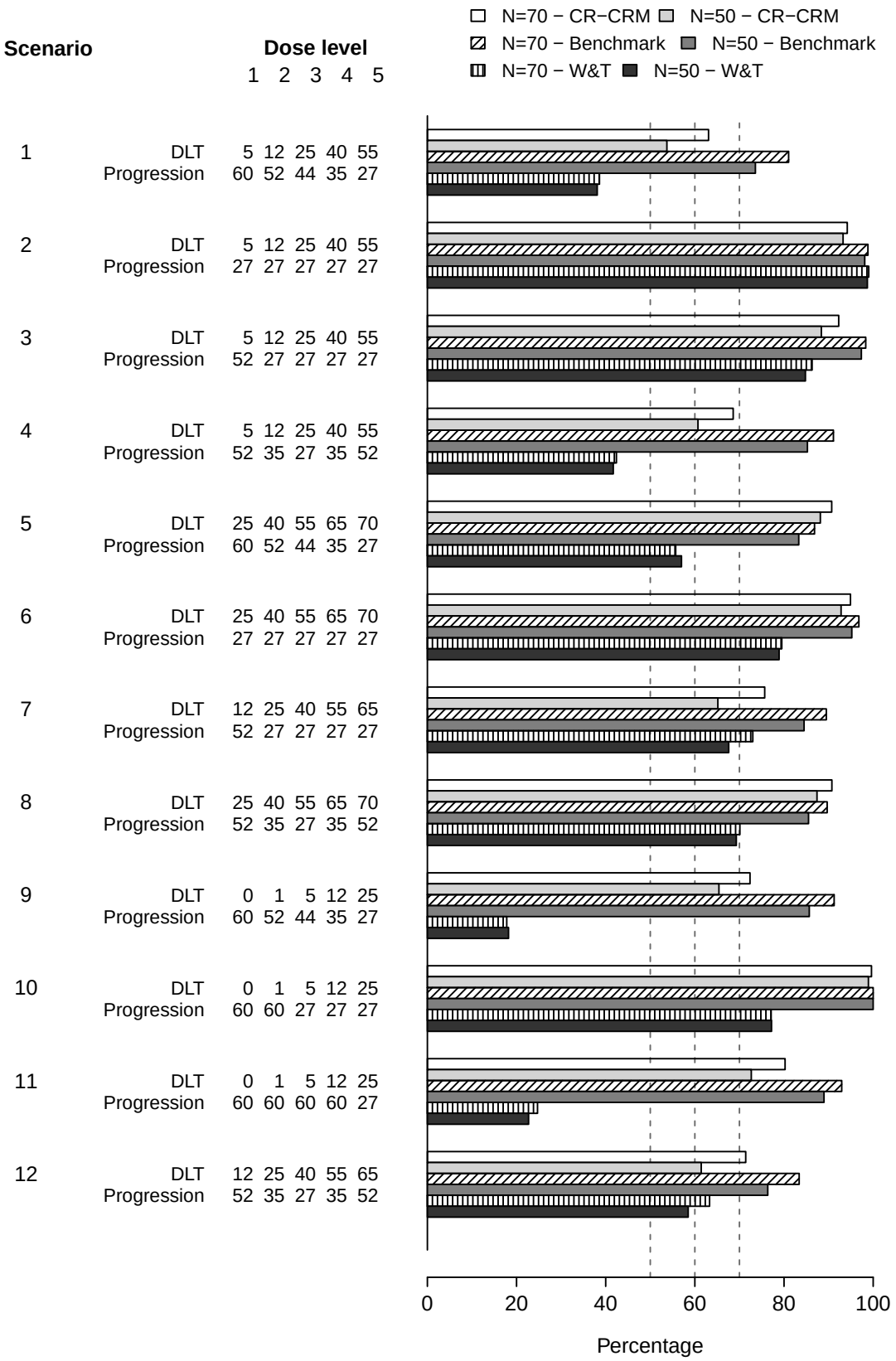

Figure S6: Probability of correctly selecting (PCS) the best dose by total sample size: $N = 35 + 35$ or $N = 25 + 25$, in Set 1 scenarios and with target maximum toxicity risk at $t^\star = 25\%$, estimated over 10000 simulations (CR-CRM and Benchmark) and 5000 simulations (Wages and Tait's design, W&T). Marginal cumulative incidences of toxicity and progression at $t^\star$ are given by scenario and dose level, in percent.

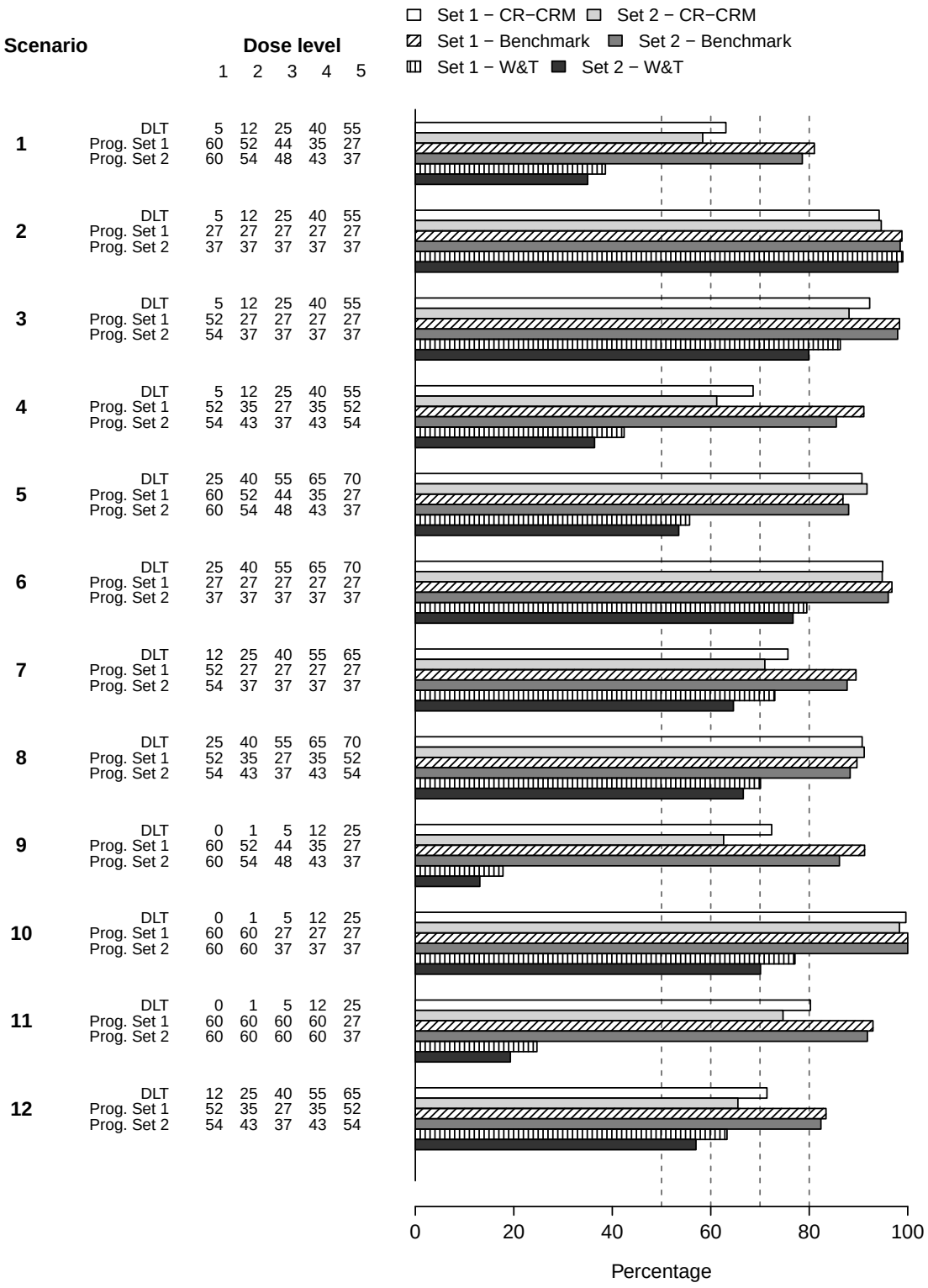

Figure S7: Probability of correctly selecting (PCS) the best dose by progression scenario, Set 1 or 2, with $N = 35+35$ and with target maximum toxicity $= 25\%$, estimated over 10000 simulations (CR-CRM and Benchmark) and 5000 simulations (Wages and Tait's design, W&T). Marginal cumulative incidences of toxicity and progression at $t^\star$ are given by scenario and dose level, in percent.

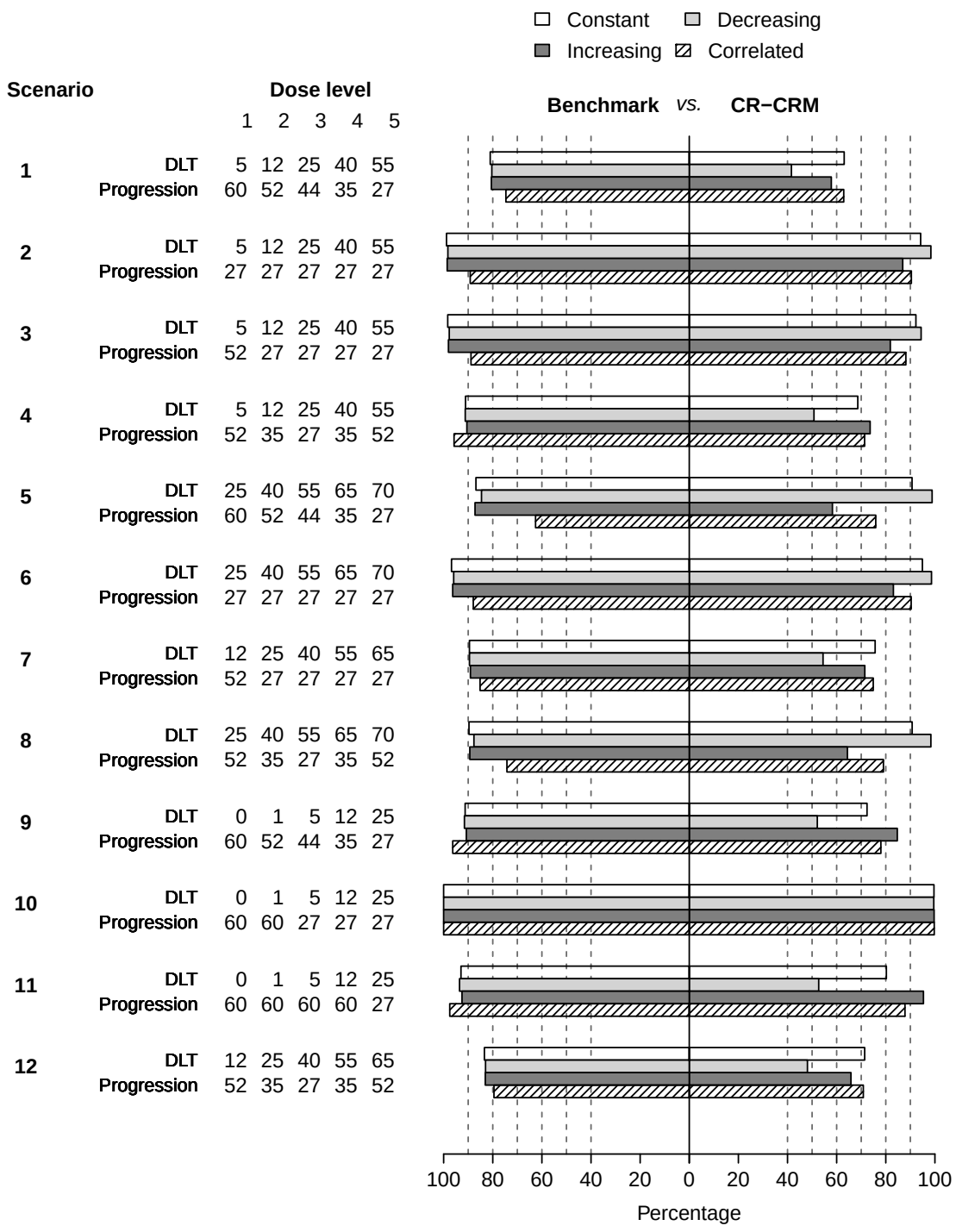

Figure S8: Sensitivity analysis: Probability of correctly selecting (PCS) the best dose by type of true hazards, in Set 1 scenarios, with $N = 35 + 35$ and with target maximum toxicity $= 25\%$, estimated over 10000 simulations. Constant: constant hazards, Decreasing: decreasing hazards, Increasing: increasing hazards, Correlated: correlated time to toxicity and time to progression using Clayton's model for data generation. Marginal cumulative incidences of toxicity and progression at $t^\star$ are given by scenario and dose level, in percent.

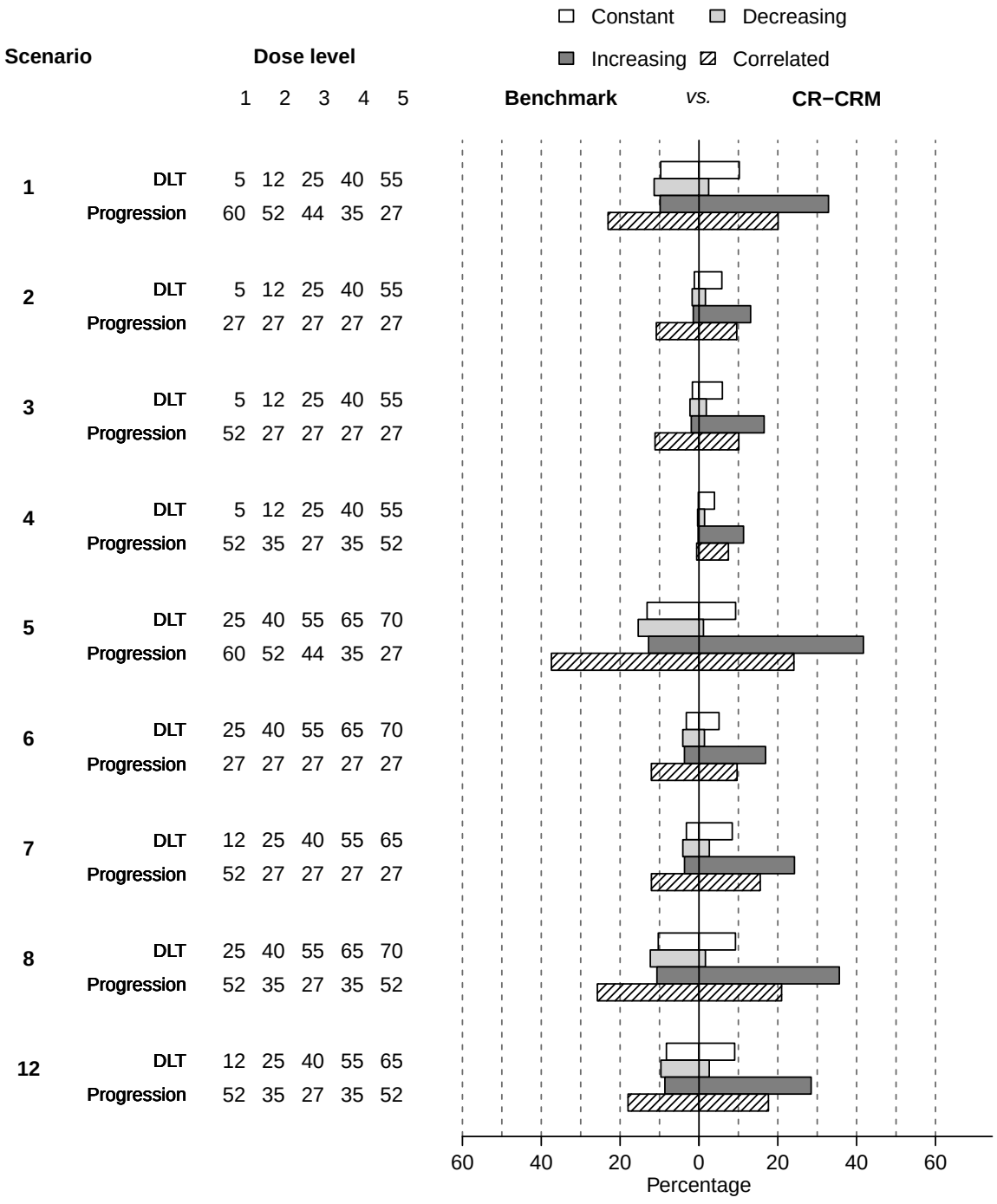

Figure S9: Sensitivity analysis: Probability of overdose selection (POS) by type of true hazards, in Set 1 scenarios, with $N = 35+35$ and with target maximum toxicity $= 25\%$, estimated over 10000 simulations. Constant: constant hazards, Decreasing: decreasing hazards, Increasing: increasing hazards, Correlated: correlated time to toxicity and time to progression using Clayton's model for data generation. Marginal cumulative incidences of toxicity and progression at $t^\star$ are given by scenario and dose level, in percent.

References

Beyersmann, J., Latouche A., Buchholz A., and Schumacher M. (2009). Simulating competing risks data in survival analysis. *Statistics in Medicine*, 28, 956–971.

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

Lee, S. M. and Cheung, Y. K. (2009). Model calibration in the continual reassessment method. *Clinical Trials*, 6, 227–238.

Mozgunov, P., Jaki T., and Paoletti X. (2020). A benchmark for dose finding studies with continuous outcomes. *Biostatistics*, 21, 189–201.

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

Yuan, Y. and Yin, G. (2009). Bayesian dose finding by jointly modelling toxicity and efficacy as time-to-event outcomes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 58, 719–736.