Supporting Web Materials for Surv-CRM-12: A Bayesian Phase I/II Survival CRM for right-censored toxicity endpoints with competing disease progression

By A. Andrillon, S. Chevret, SM. Lee, and L. Biard

1 | BENCHMARK METHOD

To obtain benchmark performances for our trial objective given a scenario of true cumulative probability of DLT and progression, the trial sample size N and the observation window t^*, we applied a nonparametric benchmark method for dose-finding trials (O'Quigley *et al.*, 2002). The approach for complex designs developed by Cheung (2014) was considered in our setting of time-to-event competing endpoints. The algorithm presented in Table S1 provides step-by-step guidance on how the benchmark can be constructed based on simulated patient data.

TABLE S1 Template to compute a benchmark for dose finding trial objective with a right-censored toxicity endpoint based on S simulated trial outcomes for a given scenario of true cumulative probability of DLT of J doses, $\{p_{1j}\}_{j=1}^{J}$, and true cumulative probability of disease progression of J doses, $\{p_{2j}\}_{j=1}^{J}$, at time $t^\star$.

Algorithm:

1. Define $F_{1j}(\cdot)$ and $F_{2j}(\cdot)$, the CDF of time-to-toxicity and time-to-progression, and specify the observation window $[0, t^*]$
2. Derive instantaneous hazards λ_{1j} and λ_{2j} consistent with $F_{1j}(\cdot)$ and $F_{2j}(\cdot)$ for each dose level according to pre-specified scenarios by finding the values of λ_{1j} and λ_{2j} for which $p_{1j} = F_{1j}(t^*, \lambda_1 j)$ and $p_{2j} = F_{2j}(t^*, \lambda_2 j)$
3. Generate a sequence of n patients' toxicity profiles $\{u_i\}_{i=1}^{n}$ from the uniform distribution $\mathbb{U}(0, 1)$
4. Obtain time-to-event using $T_{ij} = Q^{-1}(u_i, \lambda_{1j} + \lambda_{2j})$ for each patient i and each dose level j, with $Q(.)$ the cumulative distribution function of the exponential distribution
5. Determine the type to event by a random drawn from a Bernoulli distribution with probability $\frac{\lambda_{1j}}{\lambda_{1j}+\lambda_{2j}}$ for toxicity ($k = 1$; and otherwise progression, $k = 2$)
6. Apply fixed censoring at t^* to obtain trial observations : $X_i = min(T_i, t^*)$ and $Y_i = \mathbb{I}(T_i \leq t^*)$
7. From all (X_i, Y_i) compute non parametric estimates of the cumulative incidences $\hat{F}_{1j}(t^*, \lambda_{1j})$ and $\hat{F}_{2j}(t^*, \lambda_{2j})$ applying the Gray estimator for each dose level j
8. Estimate the OD, applying the OD definition to $\hat{F}_{1j}(t^*, \lambda_{1j})$ and $\hat{F}_{2j}(t^*, \lambda_{2j})$ consistently with the dose-finding objective of the evaluated design
9. Repeat steps 2-8 for $s = 1, ..., S$ simulated trials
10. Obtain a benchmark estimate of the probability of selection as OD by dose level averaging results over the S simulated trials

2 | CALIBRATION

Calibration was performed to jointly obtain the dose skeleton and the prior variance for the simulation study using the indifference intervals approach and least informative variance (Lee and Cheung, 2011). We calculated $\sigma_{\beta_1}^{LI}$ for each value of δ and selected the $(\delta, \sigma_{\beta_1}^{LI})$ that maximizes the average PCS across the calibration set. More specifically, we performed simulations using the Surv-CRM-12 under a set of calibration scenarios, such that the true probabilities of DLT and progression follow the plateau configuration, where $\mu_{1j} = F_{1L}$ and $\mu_{2j} = F_{2U}$ for $j < l$, $\mu_{1j} = F_{1U}$ and $\mu_{2j} = F_{2L}$ for $j > l$ and $\mu_{1j} = \theta_{DLT}$ and $\mu_{2j} = \theta_{PROG}$

for $j = l$ where $l = 1, ..., J$, $F_{1L} = \theta_{DLT}/(2 - \theta_{DLT})$, $F_{1U} = 2\theta_{DLT}/(1 + \theta_{DLT})$, $F_{2U} = 2\theta_{PROG}/(1 + \theta_{PROG})$ and $F_{2L} = \theta_{PROG}/(2-\theta_{PROG})$. As a sensitivity analysis, calibration was also performed to obtain the dose skeleton for the simulation study only using the indifference intervals approach (Lee and Cheung, 2009) and given a normal prior standard deviation fixed at $\sqrt{1.34}$, i.e. the value commonly used for CRM (O'Quigley and Shen, 1996). For the latest calibration approach, based on 2,000 simulations, we set δ at 0.06, as providing the highest average PCS across the calibration scenarios. As illustration, figure S1 reported the average PCS by interval half-width, across calibration scenarios, over 2000 simulations according the calibration approach. Table S2 presents simulation results using the Surv-CRM-12 with dose skeleton calibration but without performing calibration on the prior variance.

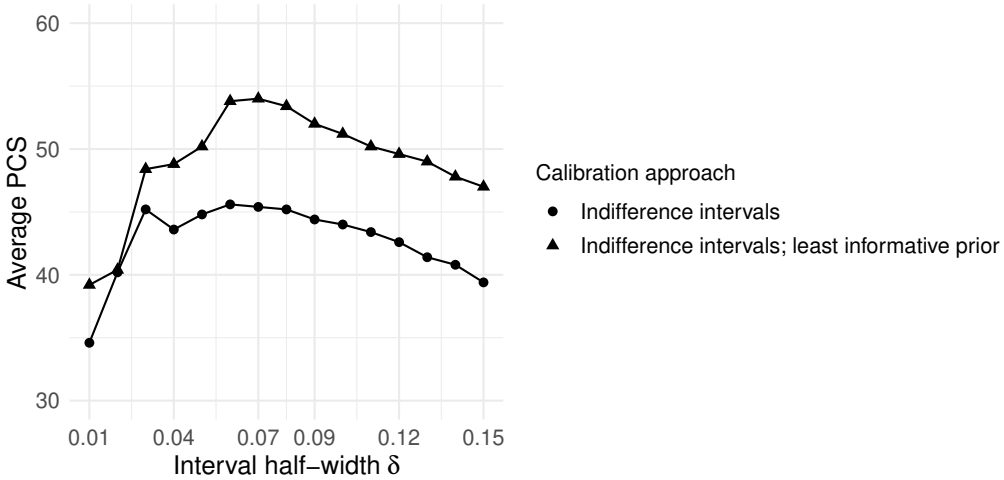

FIGURE S1 Average probability of correct selection (PCS) of the optimal dose across the set of calibration scenarios, by calibration interval half-width according to the calibration approach, with N = 45.

TABLE S2 Simulation results for Sc1 to Sc12 of the Surv-CRM-12 design; percent of stopped trials for safety (P_{stop}); percent of selection (PS), number of overdose (No. OV); number of observed DLT (No. DLT); number of observed progression (No. Prog) and number of patients treated with the true OD (No. OD) during the trial. 10,000 simulated trials with $N = 45$, $\pi_{DLT} = 0.25$, $\epsilon_P = 0.1$, and $\sigma_{\beta_1} = \sigma_{\beta_2} = \sqrt{1.34}$. Correct selection results based on ODs are given in boldface. (n/a: not applicable).

		PS by dose level (%)								
Scenario	**P_{stop} (%)**	**1**	**2**	**3**	**4**	**5**	**No. OV**	**No. DLT**	**No. Prog**	**No. OD**
Sc1	1	0	22	**68**	10	0	8.49	10.90	20.45	20.95
Sc2	0	0	21	**67**	12	0	7.61	10.30	9.74	18.92
Sc3	1	0	21	**67**	11	0	9.31	11.12	19.56	20.77
Sc4	0	1	1	3	**67**	28	13.39	10.02	13.45	20.13
Sc5	0	0	0	1	24	**75**	0.00	8.34	13.78	24.03
Sc6	1	0	**21**	**73**	5	0	6.60	10.45	20.39	34.13
Sc7	0	0	**1**	**29**	**65**	5	6.38	9.70	20.08	36.42
Sc8	0	0	1	**25**	66	8	25.98	9.74	16.05	12.85
Sc9	4	0	11	**41**	33	12	21.04	11.46	20.01	14.54
Sc10	3	0	1	10	**34**	53	19.65	8.99	26.28	13.94
Sc11	0	0	4	**68**	27	0	14.49	11.01	14.12	21.68
Sc12	18	**59**	22	1	0	0	18.27	12.70	21.52	26.73

3 | SENSITIVITY SIMULATIONS

3.1 | Simulation with samples sizes variations

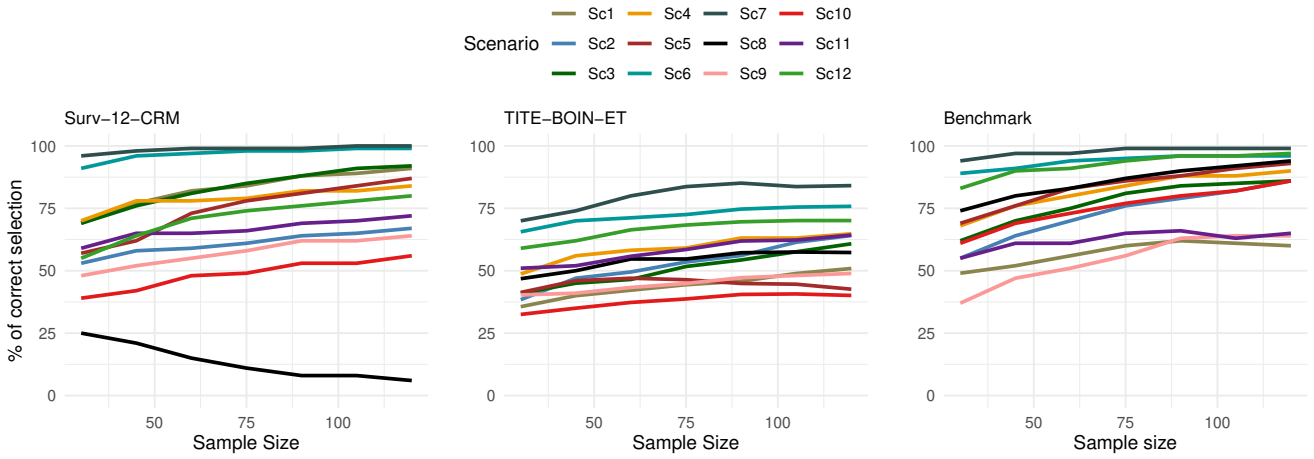

FIGURE S2 Percent of correct selection (PCS) with the Surv-CRM-12, the TITE-BOIN-ET designs and the benchmark according to the total sample size N.

Table S3 reports simulation results with a total sample size of $N = 45$ and $N = 90$ depending on the number of patients in the optional toxicity centered stage, with the proposed Surv-CRM-12 design.

TABLE S3 Simulation results for Sc1 to Sc12 of the Surv-CRM-12 design according the sample size N, and the number of patients allocated to the first optional stage, rN; percent of stopped trials for safety (P_{stop}); percent of selection (PS), number of overdose (No. OV); number of observed DLT (No. DLT); number of observed progression (No. Prog) and number of patients treated with the true OD (No. OD) during the trial. 10,000 simulated trials, $\pi_{DLT} = 0.25$ and $\epsilon_P = 0.1$. (n/a: not applicable).

		PS of selection by dose level (%)								
	P_{stop} (%)	1	2	3	4	5	No. OV	No. DLT	No. Prog	No. OD
Scenario 1										
True F_1		0.08	0.14	**0.25**	0.40	0.57				
True F_2		0.50	0.48	**0.45**	0.43	0.40				
$N = 45; rN = 15$	0	0	18	73	9	0	7.09	10.97	20.35	25.40
$N = 45; rN = 23$	0	0	18	73	9	0	7.06	10.97	20.35	25.48
$N = 90; rN = 20$	0	0	11	85	4	0	10.32	21.99	40.68	61.21
$N = 90; rN = 45$	0	0	11	85	4	0	10.19	21.99	40.65	61.29
Scenario 2										
True F_1		0.08	0.14	**0.25**	0.40,	0.57				
True F_2		0.30	0.25	**0.20**	0.15	0.10				
$N = 45; rN = 15$	0	0	17	72	10	0	7.67	11.02	9.31	25.01
$N = 45; rN = 23$	0	0	17	72	10	0	7.70	11.02	9.30	24.97
$N = 90; rN = 20$	0	0	10	85	5	0	10.97	22.08	18.48	60.65
$N = 90; rN = 45$	0	0	10	86	4	0	10.98	22.08	18.48	60.66
Scenario 3										

True F_1		0.08	0.14	**0.25**	0.40	0.57				
True F_2		0.55	0.49	**0.43**	0.36	0.30				
$N = 45; rN = 15$	0	0	17	73	10	0	7.56	11.07	19.44	25.28
$N = 45; rN = 23$	0	0	17	73	10	0	7.54	11.05	19.46	25.35
$N = 90; rN = 20$	0	0	10	85	4	0	10.98	22.15	38.68	61.17
$N = 90; rN = 45$	0	0	10	85	4	0	10.99	22.15	38.69	61.17
Scenario 4										
True F_1		0.05	0.10	0.15	**0.20**	0.25				
True F_2		0.95	0.75	0.50	**0.40**	0.30				
$N = 45; rN = 15$	0	0	0	4	79	18	7.85	9.20	13.66	26.71
$N = 45; rN = 23$	0	0	0	4	78	18	7.89	9.20	13.67	26.65
$N = 90; rN = 20$	0	0	0	0	87	13	14.90	19.52	26.47	63.79
$N = 90; rN = 45$	0	0	0	0	87	13	14.97	19.52	26.48	63.72
Scenario 5										
True F_1		0.02	0.04	0.08	0.14	**0.25**				
True F_2		0.55	0.48	0.40	0.32	**0.25**				
$N = 45; rN = 15$	0	0	0	1	34	65	0.00	7.72	14.17	18.71
$N = 45; rN = 23$	0	0	0	1	34	65	0.00	7.72	14.18	18.68
$N = 90; rN = 20$	0	0	0	0	21	79	0.00	17.63	26.34	50.95
$N = 90; rN = 45$	0	0	0	0	21	79	0.00	17.62	26.35	50.97
Scenario 6										
True F_1		0.08	**0.14**	**0.25**	0.40	0.57				
True F_2		0.50	**0.45**	**0.45**	0.45	0.40				
$N = 45; rN = 15$	0	0	19	72	9	0	6.97	10.93	20.26	36.36
$N = 45; rN = 23$	0	0	19	73	8	0	6.87	10.94	20.24	36.46
$N = 90; rN = 20$	0	0	11	85	4	0	9.99	21.91	40.51	78.35
$N = 90; rN = 45$	0	0	11	85	4	0	9.99	21.92	40.51	78.35
Scenario 7										
True F_1		0.04	**0.08**	**0.14**	**0.25**	0.40				
True F_2		0.50	**0.45**	**0.45**	**0.45**	0.40				
$N = 45; rN = 15$	0	0	0	26	67	7	3.80	9.40	20.14	40.07
$N = 45; rN = 23$	0	0	0	26	68	7	3.81	9.40	20.14	40.06
$N = 90; rN = 20$	0	0	0	15	82	3	6.09	19.96	40.23	82.78
$N = 90; rN = 45$	0	0	0	15	82	3	6.20	19.97	40.22	82.67
Scenario 8										
True F_1		0.04	0.08	**0.14**	0.25	0.40				
True F_2		0.50	0.35	**0.27**	0.35	0.50				
$N = 45; rN = 15$	0	0	0	26	68	6	25.34	9.34	15.15	15.64
$N = 45; rN = 23$	0	0	0	26	68	5	25.36	9.33	15.16	15.61
$N = 90; rN = 20$	0	0	0	16	82	2	60.71	19.77	30.30	25.14
$N = 90; rN = 45$	0	0	0	16	82	2	60.85	19.77	30.31	25.00
Scenario 9										
True F_1		0.05	0.15	**0.25**	0.30	0.35				
True F_2		0.95	0.65	**0.45**	0.35	0.25				
$N = 45; rN = 15$	0	0	14	55	26	5	14.06	10.94	20.99	20.81
$N = 45; rN = 23$	0	0	14	55	26	5	14.07	10.95	20.99	20.75

$N = 90; rN = 20$	0	0	8	64	25	2	27.66	22.43	40.61	47.55
$N = 90; rN = 45$	0	0	8	64	26	2	27.63	22.43	40.62	47.63
Scenario 10										
True F_1		0.05	0.10	0.15	**0.20**	0.25				
True F_2		0.95	0.80	0.65	**0.50**	0.55				
$N = 45; rN = 15$	0	0	1	17	48	35	10.93	8.51	26.19	18.28
$N = 45; rN = 23$	0	0	1	17	48	35	10.91	8.50	26.20	18.33
$N = 90; rN = 20$	0	0	0	8	50	42	28.72	18.08	50.34	40.39
$N = 90; rN = 45$	0	0	0	8	50	42	28.77	18.08	50.33	40.38
Scenario 11										
True F_1		0.05	0.08	**0.22**	0.35	0.55				
True F_2		0.70	0.32	**0.30**	0.28	0.26				
$N = 45; rN = 15$	0	0	4	**71**	25	0	12.64	10.75	13.78	25.74
$N = 45; rN = 23$	0	0	4	**71**	25	0	12.68	10.73	13.79	25.68
$N = 90; rN = 20$	0	0	1	**80**	19	0	23.01	21.83	27.10	59.53
$N = 90; rN = 45$	0	0	1	**81**	19	0	22.97	21.83	27.09	59.57
Scenario 12										
True F_1		**0.20**	0.40	0.50	0.55	0.62				
True F_2		**0.50**	0.45	0.40	0.38	0.36				
$N = 45; rN = 15$	4	**66**	29	0	0	0	22.18	13.72	21.22	22.82
$N = 45; rN = 23$	4	**66**	30	0	0	0	22.13	13.71	21.22	22.87
$N = 90; rN = 20$	6	**77**	17	0	0	0	33.19	24.80	43.13	56.81
$N = 90; rN = 45$	6	**77**	17	0	0	0	33.28	24.80	43.15	56.72

3.2 | Simulation of correlated time-to-toxicity and time-to-progression

To take into account the correlation between dose-limiting toxicity (DLT) and efficacy, we modeled times to DLT and progression using the Clayton model, as proposed by Yuan and Yin (2009). We thus generated correlated pairs of time-to-toxicity and time-to-progression (t_{DLT}, t_{PROG}), for each patient. The joint density of the survival times of both outcomes, $f(t_{DLT}, t_{PROG})$, was defined as:

$$f(t_{DLT}, t_{PROG}) = \frac{c+1}{c}\{S_{DLT}(t_{DLT})^{-1/c} + S_{PROG}(t_{PROG})^{-1/c} - 1\}^{-c-2} f_{DLT}(t_{DLT}) f_{PROG}(t_{PROG})\{S_{DLT}(t_{DLT}) S_{PROG}(t_{PROG})\}^{-1/c-1}$$

with $f_{DLT}(t_{DLT})$ and $f_{PROG}(t_{PROG})$ the marginal density functions for (t_{DLT}, t_{PROG}), $S_{DLT}(t_{DLT})$ and $S_{PROG}(t_{PROG})$ the survival functions for t_{DLT} and t_{PROG} respectively, and c the correlation between times to toxicity and progression.

The bivariate random variable (t_{DLT}, t_{PROG}) was generated by first simulating t_{DLT} from its marginal distribution function $f_{DLT}(t_{DLT})$, then generating t_{PROG} from its conditional distribution $f_{PROG|DLT}(t_{PROG}|t_{DLT})$. In particular, we generated two independent random profiles (u_1, u_2) from the uniform distribution $\mathbb{U}(0, 1)$. By inverse transform sampling, we then obtained $t_{DLT} = S^{-1}_{DLT}(1-u_1)$ and $t_{PROG} = S^{-1}_{PROG|DLT}(1-u_2)$. We assumed exponential marginal survival functions, S_{DLT} and S_{PROG}.

Figure S3 reports the percents of correct selection and the percents of overdose selection (POS) depending on the correlation between time-to-toxicity and time-to-progression, with the proposed Surv-CRM-12 and corresponding benchmark.

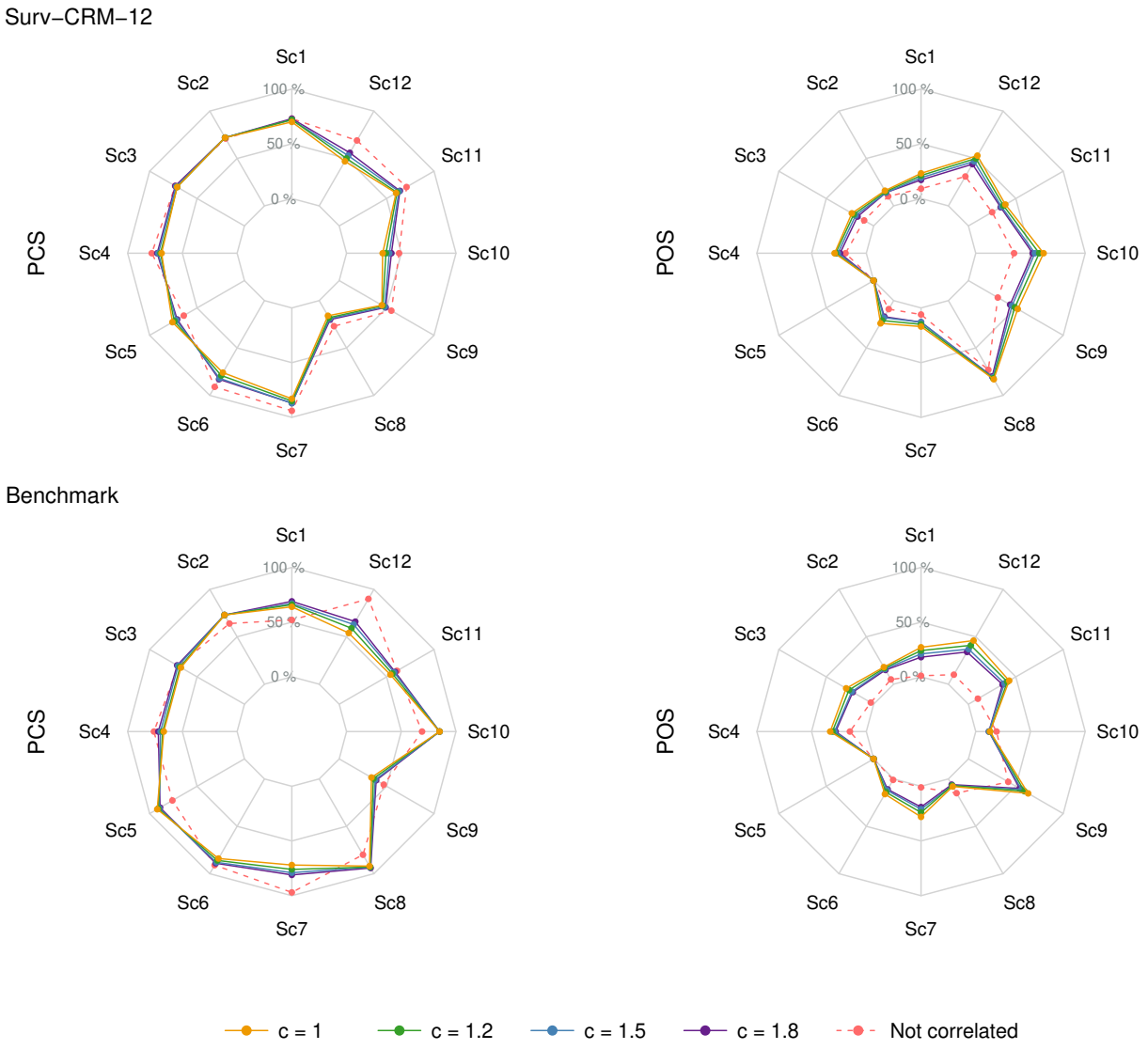

FIGURE S3 Percent of correct selection (PCS) and Percent of overdose selection (POS) with the Surv-CRM-12 and the benchmark using Clayton model for data generation according to according to the correlation between time-to-toxicity and time-to-progression. A small value of c represents a high correlation. When $c \to 0$, the correlation approaches 1 and, when $c \to \infty$, the correlation converges to 0.