ARTICLE TYPE

Supplementary Materials for Practical Implementation of the Partial Ordering Continual Reassessment Method in a Phase I Combination-Schedule Dose-Finding Trial

Pavel Mozgunov*1 | Thomas Jaki1,2 | Ioannis Gounaris3 | Thomas Goddemeier4 | Anja Victor4 | Marianna Grinberg4

1MRC Biostatistics Unit, University of Cambridge, Cambridge, UK
2Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
3Merck Serono Ltd., Feltham, UK, an affiliate of Merck KGaA, Feltham, UK
4Biostatistics, Epidemiology & Medical Writing, Merck Healthcare KGaA, Darmstadt, Germany

Correspondence
*Pavel Mozgunov, Email: pavel.mozgunov@mrc-bsu.cam.ac.uk

Present Address
MRC Biostatistics Unit, East Forvie Building, University of Cambridge, Cambridge CB2 0SR, United Kingdom.

Summary

KEYWORDS:
Combination Study, Dose-Finding, Dose-Schedule, Partial Ordering

SUPPLEMENTARY MATERIALS

Selected Orderings

1. 1, 2, 3, 6, 4, 5, 7, 8, 9, 10, 11, 12, 13, 16, 14, 15, 17, 18, 19, 20;
2. 1, 2, 3, 6, 7, 4, 5, 8, 9, 10, 11, 12, 13, 16, 17, 14, 15, 18, 19, 20;
3. 1, 6, 2, 3, 4, 7, 5, 8, 9, 10, 11, 12, 16, 13, 14, 17, 15, 18, 19, 20;
4. 1, 2, 3, 7, 4, 5, 8, 9, 10, 11, 16, 12, 13, 17, 14, 15, 18, 19, 20;
5. 1, 2, 6, 3, 7, 4, 8, 5, 9, 10, 11, 12, 16, 13, 17, 14, 18, 15, 19, 20;
6. 1, 2, 3, 6, 4, 7, 11, 12, 13, 16, 14, 17, 5, 8, 9, 10, 15, 18, 19, 20;
7. 1, 2, 3, 7, 4, 11, 12, 16, 13, 17, 14, 5, 8, 9, 10, 15, 18, 19, 20;
8. 1, 2, 6, 3, 4, 7, 11, 12, 16, 13, 14, 17, 5, 8, 9, 10, 15, 18, 19, 20;
9. 1, 2, 6, 11, 3, 4, 7, 12, 13, 16, 5, 8, 14, 17, 9, 15, 18, 10, 19, 20;
10. 1, 2, 6, 11, 3, 16, 12, 13, 7, 4, 5, 8, 14, 17, 18, 15, 9, 10, 19, 20;
Prior Probabilities of Orderings

The a-priori probabilities of each of the ten orderings specified above that correspond to the pairwise probabilities reported in Table 3 are given by

(0.12, 0.16, 0.16, 0.09, 0.16, 0.07, 0.09, 0.07, 0.04, 0.04).
TABLE S1 Operating characteristics for $n = 42, 48, 54$ following (i) the allocation according to the criterion (Next Best), (ii) the randomisation “Within Admissible Set”. “M” corresponds to the mean proportion, and $P$ corresponds to the proportion of scenarios with the percentage of optimal selections being above 40%, correct selection being above 60%, and overly toxic selection being below 25%. Results are based on 2000 simulations.
TABLE S2  Operating characteristics of the POCRM design with model averaging for \( n = 42, 48, 54 \) following (i) the allocation according to the criterion (Next Best), (ii) the randomisation “Within Admissible Set”. “M” corresponds to the mean proportion, and \( \mathbb{P} \) corresponds to the proportion of scenarios with the percentage of optimal selections being above 40%, correct selection being above 60%, and overly toxic selection being below 25%. Results are based on 2000 simulations.
Example of model recommendation output

As mentioned in the main body of the manuscript, at the planning stage of the trial, it is crucial to ensure that the proposed model leads to intuitive escalation/de-escalation decision throughout the trial. One way to ensure this is through the decision-tree presented in the main manuscript. In addition, we also provided the complete output of the model that will be shared with the trial team during the conduct of the trial. This provides a more detailed look at why the proposed model recommends the regimens as per the decision-tree but also to ensure that the output is relevant to the team to ensure efficient decision-making.

There are two types of output that were provided to the team: (i) the complete table of the estimated characteristics for all twenty regimens, and (ii) a figure with the probabilities of each regimen being overtoxic, undertoxic, and in the target toxicity range.

The detailed table were agreed to include the following information for each regimen:

- Number of patients and DLTs;
- Mean estimated toxicity risk;
- 95% Credible interval
- Probability of overdosing: \( P(\pi_r > 0.35) \)
- Probability being in the target range: \( P(0.20 \leq \pi_r \leq 0.35) \)

We have also used the following colour-coding to aid the communication of the output table:

The following colour-coding is used:

- If the overdose probability is above 25%, the corresponding cell is red.
- If the overdose probability is below 25%, there are two possibilities, the cell can be green or blue with the convention that all of these are safe according to the model but green are preferred for the escalation as follows:
  - Safe combinations corresponding to more than doubling of the total average and single dose are blue.
  - If the combination is escalated [within the chosen most likely ordering], the escalated combinations are green.
  - If the combinations is de-escalated and there are more than six safe combinations, the combinations that correspond to less than halving of the total average dose of M or 200mg dose of N are preferred (and hence are in green).

Additionally, we have provided a figure displaying for each regimen being overtotoxic (red, \( P(\pi_r > 0.35) \)), being in the target range (green, \( P(0.20 \leq \pi_r \leq 0.35) \)), and being undertoxic (blue, \( P(\pi_r \leq 0.20) \)). This figure uses the same combination-schedule grid given in Table 2 of the main body of the manuscript. An example of such figure after 1 DLT was observed for the first cohort of 3 patients at the starting combination (18) is given in Figure S1. The threshold of 75% is provided to illustrated which of the regimens are deemed overly toxic (using the probability threshold of 25%) and cannot be selected.
TABLE S3 Example of the output provided to the SRC if the first cohort of 3 patients experienced no DLTs
TABLE S4 Example of the output provided to the SRC if the first cohort of 3 patients experienced 1 DLT
### TABLE S5 Example of the output provided to the SRC if the first cohort of 3 patients experienced 2 DLTs

| 200N, 30M, S1 | 200N, 60M, S1 | 200N, 90M, S1 | 200N, 130M, S1 | 200N, 180M, S1 |
|----------------|----------------|----------------|----------------|----------------|
| **n=0** | **n=0** | **n=0** | **n=0** | **n=0** |
| DLTs=0 | DLTs=0 | DLTs=0 | DLTs=0 | DLTs=0 |

Mean Tox: 0.32  
95%CI:0.01(0.07)  
Overdose: 42.4%  
Targets: 2.3%  

| 200N, 30M, S2 | 200N, 60M, S2 | 200N, 90M, S2 | 200N, 130M, S2 | 200N, 180M, S2 |
|----------------|----------------|----------------|----------------|----------------|
| **n=0** | **n=0** | **n=0** | **n=0** | **n=0** |
| DLTs=0 | DLTs=0 | DLTs=0 | DLTs=0 | DLTs=0 |

Mean Tox: 0.24  
95%CI:0.01(0.06)  
Overdose: 25.9%  
Targets: 2.3%  

| 100N, 60M, S1 | 100N, 90M, S1 | 100N, 130M, S1 | 100N, 180M, S1 |
|----------------|----------------|----------------|----------------|
| **n=0** | **n=0** | **n=0** | **n=0** |
| DLTs=0 | DLTs=0 | DLTs=0 | DLTs=0 |

Mean Tox: 0.04  
95%CI:0.01(0.04)  
Overdose: 3.4%  
Targets: 1.3%  

| 100N, 30M, S2 | 100N, 60M, S2 | 100N, 90M, S2 | 100N, 130M, S2 | 100N, 180M, S2 |
|----------------|----------------|----------------|----------------|----------------|
| **n=0** | **n=0** | **n=0** | **n=0** | **n=0** |
| DLTs=0 | DLTs=0 | DLTs=0 | DLTs=0 | DLTs=0 |

Mean Tox: 0  
95%CI:0(0.1)  
Overdose: 0.2%  
Targets: 0%
FIGURE S1 Example of the visualisation of the probability of risk of toxicity for each combination-schedule being in the underdosing, target, and overdosing interval used the grid representation of the regimens. The example is displayed after 1 DLT was observed for the first cohort of 3 patients at the starting combination (18).