

White Paper

A Comparison of the CRM vs 3+3 in an Oncology phase 1 setting

Tom Parke

Head of Clinical Trial Solutions, Tessella plc

A Comparison of the CRM vs 3+3 in an Oncology phase 1 setting

The principal aim of an Oncology phase 1 trial is the identification of the Maximum Tolerated Dose (MTD). The majority of oncology phase 1 trials are run using a design called “3+3”[1]. Under this design, subjects are treated in cohorts of 3, and based on the number of dose limiting toxicities seen in that cohort, decisions on which dose to give the next cohort and whether to stop the trial are made.

In [2], Reiner et al argue that this design does not give very good operating characteristics. “the probability of recommending the [correct] MTD at the end of the trial ... never exceeds 44% and is most often closer to 30%”. Despite this, the 3+3 design remains in common usage due to its simplicity, the straightforward (and appealing) nature of its decision rules and its familiarity.

Possibly one other reason is the difficulty of doing a lot better. In this essay we look at using the oft touted alternative to the 3+3, the Continuous Reassessment Method (CRM) proposed by O’Quigley in 1990 [3] and given some modest but important design tweaks by Goodman in 1995 [4].

Comparison

In [2] a set of 10 scenarios are given, for the possible probability of toxicity on each of 6 doses. The doses are in increasing strength and hence in all scenarios there is an increasing probability of toxicity.

| Scenarios | Doses | | | | | |
|-----------|-------------|-------------|-------------|-------------|-------------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| S1 | 0.09 | 0.15 | 0.25 | 0.38 | 0.55 | 0.65 |
| S2 | 0.02 | 0.09 | 0.16 | 0.27 | 0.35 | 0.45 |
| S3 | 0.01 | 0.05 | 0.08 | 0.2 | 0.35 | 0.5 |
| S4 | 0.05 | 0.1 | 0.22 | 0.32 | 0.4 | 0.5 |
| S5 | 0.24 | 0.39 | 0.5 | 0.6 | 0.65 | 0.7 |
| S6 | 0.05 | 0.1 | 0.25 | 0.35 | 0.5 | 0.7 |
| S7 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.75 |
| S8 | 0.15 | 0.26 | 0.4 | 0.57 | 0.65 | 0.7 |
| S9 | 0.06 | 0.09 | 0.12 | 0.24 | 0.4 | 0.6 |
| S10 | 0 | 0.01 | 0.04 | 0.09 | 0.24 | 0.49 |

The figures in bold show the dose that it is desired the phase 1 trial identify as the MTD (in Scenario S7 it is actually required that no dose is identified as the MTD). These are consistent with the 3+3 design as the mean toxicity of the doses 3+3 selects is ~25%.

The results of simulating trials using a 3+3 design with these scenarios will be compared with a CRM based design. Both designs will use cohorts of 3 subjects at a time and start allocating at the lowest dose. The CRM will target the highest dose with an estimated toxicity of 25% or less.

The CRM’s escalation will be limited to incrementing no more than one dose strength above the highest dose to which a cohort has already been allocated. One variant of the 3+3 is to include a ‘validation’ cohort. Once the 3+3 stopping rule applies, if the dose level selected as MTD has only one cohort tested at that dose level, then a second cohort is tested at that level. If 2 or more toxicities are observed in the validation cohort then the selection of MTD drops a level and the process repeated. This reduces the probability that the 3+3 selects too high a dose for MTD.

The CRM stops when either the maximum number of cohorts is reached, or the probability that one dose is the MTD exceeds some chosen threshold. To give a similar sample size to the 3+3 over these scenarios, a threshold of 60% is used.

We present 3 metrics estimated from 1,000 simulations of each design under each scenario – the average sample size, the probability of selecting the MTD and the probability of selecting a dose above the ‘desired MTD’.

| Scenario | 3+3 | | | CRM | | |
|----------|------------------|--------------------------|------------------------|------------------|--------------------------|------------------------|
| | Avg. Sample Size | Ppn correct MTD selected | Ppn above MTD selected | Avg. Sample Size | Ppn correct MTD selected | Ppn above MTD selected |
| S1 | 15.4 | 29.7% | 11.1% | 15.9 | 40.9% | 13.1% |
| S2 | 18.5 | 34.9% | 36.6% | 18.0 | 33.2% | 5.5% |
| S3 | 20.0 | 38.3% | 20.5% | 17.3 | 54.1% | 15.8% |
| S4 | 17.0 | 32.1% | 22.9% | 16.8 | 49.1% | 22.4% |
| S5 | 9.6 | 82.2% | 17.8% | 10.5 | 85.9% | 14.1% |
| S6 | 16.5 | 31.0% | 17.0% | 17.5 | 53.0% | 15.3% |
| S7 | 8.4 | 86.6% | 13.4% | 8.4 | 91.0% | 9.0% |
| S8 | 12.0 | 33.6 | 10.4% | 14.6 | 35.6% | 12.4% |
| S9 | 18.4 | 33.0% | 9.7% | 15.7 | 37.7% | 9.0% |
| S10 | 21.6 | 40.4% | 10.1% | 16.3 | 48.6% | 0.0% |

Only in scenario 2 does the 3+3 method select the correct MTD more often than the CRM (34.9% vs. 33.2%); in several scenarios (S1, S3, S4 & S6) the CRM selects the correct MTD significantly more often than the 3+3 (more than 10 percentage points difference).

In only 2 scenarios (S1 and S8) does the 3+3 select dose levels above the correct MTD with less frequency than the CRM, the differences are marginal (~2 percentage points), in all the others the CRM is better than the 3+3 in this regard as well, sometimes significantly so (S2 & S10).

The CRM has the further advantage that, unlike the 3+3 its operating characteristics can be easily optimised in light of the current circumstance, different levels of toxicity can be targeted, different cohort sizes used and different levels of accuracy required before stopping (offering better determination of the MTD at the cost of greater sample size). For example:

| Scenario | CRM stop when 1 dose 70% likely to be MTD | | | CRM, stop when 1 dose 60% likely to be MTD & 4 subjects per cohort | | |
|----------|---|--------------------------|------------------------|--|--------------------------|------------------------|
| | Avg. Sample Size | Ppn correct MTD selected | Ppn above MTD selected | Avg. Sample Size | Ppn correct MTD selected | Ppn above MTD selected |
| S1 | 22.9 | 44.6% | 8.6% | 19.6 | 44.2% | 8.9% |
| S2 | 25.5 | 34.6% | 3.8% | 20.8 | 30.7% | 6.1% |
| S3 | 23.8 | 67.8% | 9.9% | 19.8 | 52.6% | 15.2% |
| S4 | 25.1 | 59.5% | 15.6% | 20.4 | 55.4% | 19.6% |
| S5 | 15.3 | 90.0% | 10.0% | 15.2 | 84.6% | 15.4% |
| S6 | 25.3 | 57.9% | 15.1% | 21.2 | 59.0% | 13.9% |
| S7 | 10.3 | 95.7% | 4.3% | 12.2 | 90.0% | 10.0% |
| S8 | 20.3 | 36.0% | 9.1% | 19.2 | 43.3% | 13.4% |
| S9 | 22.3 | 41.5% | 4.6% | 19.7 | 34.5% | 5.0% |
| S10 | 22.1 | 43.3% | 0.0% | 10.1 | 50.0% | 0.0% |

These show varying degrees of improvement in the estimation of the 'true' MTD, for the cost of 3-6 additional subjects. Interestingly in scenario 2 no significant improvement is seen – but this can be traced to the CRM target rule and the scenario. In scenario 2 the simulated probability at the target dose 4 is 27%, and the target rule we have used for the CRM is the highest dose with probability of toxicity of less than 25%, so these designs are increasing the probability of selecting the dose below this one.

Finally – your mileage may vary. The relative merits of the 3+3 and the CRM are dependent of the toxicity profile of the drug being studied.

References

- [1] Korn E, Midthune D, Chen T, et al: A Comparison of Two Phase I Trial Designs, *Statistics in Medicine* 13: 1799-1806, 1994
- [2] Reiner E, Paoletti X, O'Quigley J: Operating Characteristics of the Standard Phase 1 Clinical Trial Design, *Computational Statistics & Data Analysis* 303-315, 1999
- [3] O'Quigley J, Pepe, M, Fisher, L: Continual Reassessment Method: A Practical Design for Phase I Clinical Trials in Cancer, *Biometrics* 46: 33-48, 1990
- [4] Goodman S, Zahurak M, Piantadosi S: Some Practical Improvements in the Continual Reassessment Method for Phase I Studies, *Statistics in Medicine* 14: 1149-1161, 1995

Tessella plc 26 The Quadrant, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YS, UK
T: +44 (0)1235 555511 | F: +44 (0)1235 553301 | E: info@tessella.com

Tessella Inc 233 Needham Street, Suite 300, Newton, MA 02464, USA
T: 1 617 454 1220 | E: info@tessella.com



Tessella – successfully delivering IT and consulting services to world leaders in R&D, science and engineering.

For decades, Tessella has been successfully delivering IT and consulting services to world leaders in R&D, science, and engineering. Through the application of scientific methods and rigorous quality procedures, we enable clients in life sciences, energy, the public sector, and consumer industries to achieve a wide range of objectives, including, forecasting floods, developing fusion power, enhancing military sensor capability, improving drug discovery and development efficiency, and reducing risk to health and the environment in the extraction and production of oil and gas. With offices in Europe and North America, global companies rely on Tessella for business critical assignments.

Copyright © Tessella plc 2010, all trademarks acknowledged. Issue: V1.R0.M0 | Apr-10

www.tessella.com

