GOAL OF PHASE 1 STUDY IN ONCOLOGY

Primary goals
When designing a clinical trial, the goals must be clearly and specifically defined, and this should guide decision making on the details of the study. The goals of phase 1 oncology studies may include: 1) determine maximal tolerated dose (MTD), 2) provide recommended phase 2 dose (RP2D), 3) describe/characterize dose-limiting toxicity (DLT), 4) expand understanding of the investigational product, e.g. pharmacokinetics, efficacy. These goals may be coupled with business/development timeline (e.g., as quickly as possible) and/or cost goals (with fewest possible subjects). For novel targets or mode of action, these many goals are further complicated by limited information about clinical efficacy and toxicity. The ethical dilemma regarding dose selection has evolved from being centered on limiting the number of patients who experience toxicity to also minimize underdosing of patients in need of treatment.

Translational goals
Scientific curiosity may pose other questions/goals for an oncology phase 1 trial. These may include detecting target expression or pathway activity, observing target engagement by the drug (e.g., plasma/tissue concentration), and other translational goals. Some development teams get lost discussing tissue heterogeneity in immunohistochemistry, sampling in biopsies, scanning, and analytic methodology of magnetic resonance imaging, etc. These goals, while entertaining, should be considered as nice-to-have, and probably only considered to be incorporated into the study if they do not distract from the main objectives. Answering such translational questions can be done in later stages of development or in separate studies.

Prioritized goals
To summarize, an ideal phase 1 design in oncology may be one that enables identification of the MTD/RP2D as quickly as possible with a fewest patient number with acceptable precision. Accurate determination of MTD critically depends on 2 factors: 1) definition and clinical interpretation of DLT, and 2) trial design, including dose selection and escalation. This paper will discuss the latter factor.
TRADITIONAL RULE-BASED 3+3 DESIGN

Origins of 3+3
The “Up and Down Test” or “the staircase method” or “Bruceton analysis” was described originally by Dixon and Mood in 1948 as a way of analyzing sensitivity and sensitiveness tests of explosives [1]. The analysis relies on two parameters: stimulus and step size. After a stimulus is provided to the sample of explosives, a positive result leads to a decrement of step size, while a negative result leads to increased stimulus. The test continues with each sample tested at a stimulus 1 step up or down from the previous stimulus if the previous result was negative or positive.

Use of 3+3 in oncology trials
This up-and-down design was first described in oncology trials by Storer in 1989 [2]. 98.4% of 1,235 oncology phase 1 trials conducted between 1991 and 2006 used variations of 3+3 design [3].

From the explosive testing, the stimulus has changed to drug treatment, observation of the result has changed to DLT, and step has become dose. 3 patients entered into a dose; if no toxicity, then dose is escalated; if 1 patient experiences DLT, additional 3 patients are entered; if 2 patients experience DLT, then dose escalation is stopped, or dose is de-escalated.

While explosives are manufactured with reproducible methods, patients participating in oncology trials are highly heterogenous. That is, the rate of explosion is very predictable while drug toxicity in individual patient with variable physiology and pathology is much less likely to be reproducible.

MTD premise
There are a few premises about MTD in oncology trial: 1) monotonous dose-toxicity relationship, 2) dichotomous definition of toxic and non-toxic (pre-defined symptoms of pre-defined severity), 3) generally regarded limit for the incidence of DLT is 33%.

Limitations of 3+3
If the main objective of an oncology phase 1 trial is to determine the MTD/RP2D, then the limitations of 3+3 should be heavily weighed in selecting the trial design. The 3+3 design has a low probability of selecting true MTD [4] and tends to underestimate MTD. Industry experience has also shown high variability in MTD estimates, yielding different MTDs for the same compound from different studies.

The design ignores history and only looks at the last 3 patients that came in; qualitative different situations lead to the same result. For example, no DLT in 5 successive doses followed by 2 DLTs out of 6, would lead to the same result as and 1/6 DLTs for 5 successive doses followed by 2 DLT out of 3. The former scenario may lead to underestimation of MTD, whereas the last 3 subjects in the latter scenario may have been overdosed.

EVOLUTION TO CONTINUOUS REASSESSMENT METHOD (CRM)

Improvement in selecting MTD
The 3+3 design’s main limitation is low probability of selecting the true MTD while the main goal of an oncology phase 1 trial is to find the MTD. So, a design that is not optimally fit-
for-purpose was used in the majority of the studies. In attempts to improve MTD selection, the CRM was proposed in 1990 [5]. The CRM uses all available information from previous cohorts to guide dose assignment for the next cohort rather than the last 3 or 6 patients. The next recommended dose is based on posterior toxicity probability being closest to target. This method indeed did a better job of identifying the MTD.

Limitations in dose-toxicity modeling
Despite being able to better select MTD, the CRM was not sufficient in modeling the dose-toxicity curve since there are not enough parameter model with not enough parameters to describe the whole curve. U.S. Food and Drug Administration (FDA) raised concerns about three trials using CRM method in which patients were overdosed, exposed to doses with high rates of DLTs [6].

MODIFIED CRM

Bayesian logistic regression method (BLRM)
A better model that utilizes all available information (preclinical, historic studies as well as incoming data from ongoing trials), and able to reflect a wide range of possible dose-DLT relationships was sought. In 2008, Neuenschwander, a Novartis statistician introduced the BLRM [7]. For more than a decade, Novartis Oncology has conducted most of its early development trials using the BLRM, and now exclusively uses this method for phase 1.

Escalation with over-dose control (EWOC)
The key difference between CRM and BLRM is the use of EWOC to reduce safety risks. EWOC exposes patients to drug doses that are below a pre-defined toxicity probability level. \( p(d) \) = probability of having DLT in the first cycle at dose ‘d.’ \( p(d) \) is usually 25%. This means that the probability of overdosing should be less than 25%. In addition to overdose control, there is an escalation maximal, which is a certain percentage compared to already tested levels. The escalation maximal is usually 100%, i.e. the next dose cannot exceed twice that of the previous dose even if higher doses satisfy EWOC criterion.

What does Bayesian model deliver?
BLRM delivers a binding rule regarding safety and a non-binding recommendation for the next level. Based on EWOC, it provides a maximum dose for the next cohort that cannot be overruled. Based on target toxicity rate, it guides dose selection, but other data and clinical expertise are integrated, and in the end, clinical decision is made.

Frequent misunderstandings about BLRM
There are a few misunderstandings about the BLRM, mostly based on confusion with the CRM.

- **BLRM allows large dose escalations compared to 3+3.**
  It depends. Accumulated data may allow bigger dose escalations down the road, but the escalation scheme is usually similar to 3+3, especially during early cohorts.

- **BLRM does not require a pre-specified dose table like the 3+3, can select any dose based on the dose-toxicity curve.**
  No. Pre-defined dose levels for escalation should be specified in the protocol for modeling. However, it can be stated in the protocol that based on modeled toxicity probability, intermediate dose levels may be chosen. One frequent misperception about adaptive trials is that flexibility extends to unplanned modification to trial design and
details, so studies can be quickly started without much forethought. The FDA guidance on adaptive design emphasizes that “adaptations” should be prospectively planned, i.e., opportunities for modification should occur in one or more pre-specified aspects, with plans detailed in the protocol.

- **BLRM saves significant time in trial conduct.**

  BLRM probably does not save time with few doses, but may save time when more dose levels are tested. If an inexperienced trial team initiates the trial with insufficient preparation, the time needed for dose escalation may be significant and result in longer trial duration compared to 3+3.

- **The computer will decide what the next dose should be.**

  The computer calculates probabilities for toxicity for each of the pre-specified doses. It's a clinical decision what dose the next cohort will receive guided on the probabilities calculated by the model.

### BLRM VS. 3+3

**Key comparisons**

The following lists key comparative properties between BLRM and 3+3.

- Finding correct MTD: > 60% vs. ≤ 30%
- Risk of overdosing: similar to overdose control in Bayesian
- Risk of underdosing: return to 100% escalation steps easily possible vs. risk of early DLT with small escalation
- Inclusion of other data in dose-finding (e.g., Pharmacokinetic/pharmacodynamic, biomarkers): possible vs. not possible
- Sample size: flexible vs. inflexible
- Timelines: similar
- Acceptance by regulators: similar (FDA, European Medicines Agency)
- Statistical complexity (dependence on statistician): experienced clinical research organization, validated software, competent statistician needed for Bayesian vs. 3+3 is not a statistical method
- Combination dose-finding: feasible in Bayesian

As an example, if 2 DLTs were observed in a dose level, in the rule-based 3+3, the rule is to stop dose escalation. One dose level below this dose will become the MTD/RP2D. In BLRM, all accumulated data are used to model the toxicity probability curve. Based on modeling, if higher doses meet EWOC criterion, further escalation can be made. Additionally, de-escalations can be made with intermediate doses to more accurately estimate MTD.

### IMPLEMENTATION OF BLRM

**Considerations in trial preparation**

Attention should be paid to make sure that all sections of the protocol are aligned with the BLRM. A mistake that can be made by a company conducting BLRM may be that only the statistical section of the protocol details the BLRM while the other sections are remnants of the 3+3, and the protocol as a whole does not make logical sense. In addition to the statistical section, the following sections also need to be aligned with the BLRM: endpoints (definition of MTD), trial design and plan (e.g., specify that dose escalation and cohort size will be guided...
by BLRM with overdose control, and how clinical decision will be made), selection of doses (EWOC criterion, safety monitoring committee [SMC]), and safety monitoring committee.

The trial team must be educated and prepared to conduct BLRM. Continuous data cleaning, a seamless collaboration between site management, data management, project management, SMC, statistician are necessary to compress the time needed for dose escalation.

**SMC**
The SMC should oversee dose escalation. BLRM is model-guided and provides non-binding recommendations. The final decision on dose escalation should be a clinical one. The trial team cannot make this decision by themselves as only the investigator knows the qualitative aspects of the DLTs. It is recommended to have more than one investigator to have multiple opinions. The SMC will decide the next dose to be investigated (escalate, de-escalate), sample size for the next cohort, MTD/RP2D reached or not. In order to save time, it may be specified in the protocol that if there is no DLT, then escalation to the next pre-specified dose level will be made without SMC.

The SMC charter should be finalized prior to study initiation. The charter should clearly define the responsibilities of the SMC, define members of the SMC, what data outputs SMC members will receive prior to each meeting, decision making, timelines (e.g., 5 days after the last patient of a cohort finishes MTD evaluation period). The SMC planning should be worked out and simulations performed by inter-disciplinary teams, e.g., after site management cleans data, data management needs 1 day to prepare data transfer to SMC, statistician will need 1 day to provide BLRM model-guided non-binding recommendations, 3 days for SMC to review, 1 day for SMC meeting and decision making. Without sufficient planning, this process can take weeks. Insufficiently informed SMC unfamiliar with the BLRM may be unable to reach timely consensus for decisions for the next cohort.

**When not to use BLRM**
In reality, clinicians may be used to 3+3, and hesitant to skipping pre-defined doses despite meeting EWOC criterion. Training of personnel, validated software, great logistic effort (collecting, validating data for analysis after every few patients to determine optimal next dose level) may not be feasible for some sponsors. In such cases, attempting BLRM may be faced with many operational hiccups, and ultimately the losses in quality, time, cost may outweigh the advantage of better MTD estimates. All in all, the trial design should focus on meeting clearly defined objectives while considering practical aspects of trial conduct.

**REFERENCES**

1. Dixon JA, Mood AM. A method for obtaining and analyzing sensitivity data. J Am Stat Assoc 1948;43:109-126.
2. Storer BE. Design and analysis of phase I clinical trials. Biometrics 1989;45:925-937.
3. Rogatko A, Schoeneck D, Jonas W, Tighiouart M, Khuri FR, Porter A. Translation of innovative designs into phase I trials. J Clin Oncol 2007;25:4982-4986.
4. Thall PF, Lee SJ. Practical model-based dose-finding in phase I clinical trials: methods based on toxicity. Int J Gynecol Cancer 2003;13:251-261.
5. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. Biometrics 1990;46:33-48.

6. Lard G. Implementation of Bayesian Logistic Regression for dose escalation at Novartis Oncology. https://www.slideserve.com/gerry/implementation-of-bayesian-logistic-regression-for-dose-escalation-at-novartis-oncology. Accessed February 26, 2020.

7. Neuenschwander B, Branson M, Gsponer T. Critical aspects of the Bayesian approach to phase I cancer trials. Stat Med 2008;27:2420-2439.