Dose Transition Pathways:

Yap, Christina; Billingham, Lucinda J; Cheung, Ying Kuen; Craddock, Charles; O'Quigley, John

DOI:
10.1158/1078-0432.CCR-17-0582

Document Version
Peer reviewed version

Citation for published version (Harvard):
Yap, C, Billingham, LJ, Cheung, YK, Craddock, C & O'Quigley, J 2017, 'Dose Transition Pathways: The missing link between complex dose-finding designs and simple decision-making', *Clinical Cancer Research*, vol. 23, no. 24, pp. 7440-7447. https://doi.org/10.1158/1078-0432.CCR-17-0582

Title Page

Title

Dose Transition Pathways: The missing link between complex dose-finding designs and simple decision-making

Authors and affiliations

Christina Yap
Cancer Research UK Clinical Trials Unit, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, B15 2TT, UK

Lucinda J Billingham
Cancer Research UK Clinical Trials Unit, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, B15 2TT, UK

Ying Kuen Cheung
Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY 10032, U.S.A

Charlie Craddock
Centre for Clinical Haematology, Trust Headquarters, Main Drive, Queen Elizabeth Hospital, Birmingham, B15 2TH, UK

John O'Quigley
Université Pierre et Marie Curie-Paris VI, 4 place Jussieu, Paris 75005, France

Running Title: Dose Transition Pathways in model-based dose-finding designs

Keywords:
model-based designs, phase I dose-finding, rule-based designs, continual reassessment method, 3+3,

Additional Information

Funding
This work is funded by Cancer Research UK (C22436/A15958) and Bloodwise. The funders had no role in the study design, data analysis, interpretation or writing of the report. The corresponding author had full access to all the data and had final responsibility to submit for publication.

Corresponding author
Dr Christina Yap
Cancer Research UK Clinical Trials Unit
Institute of Cancer and Genomic Sciences
College of Medical and Dental Sciences
University of Birmingham
Birmingham B15 2TT
United Kingdom

Tel: +44 121 4151048
Fax: +44 0121 414 7471
✉ c.yap@bham.ac.uk

Conflicts of interest
Dr. Yap reports personal fees from Celgene (Honoraria), outside the submitted work; Prof. Craddock reports personal fees from Celgene (Honoraria), personal fees from Celgene (Speaker Panel), grants from Celgene (Research funding), outside the submitted work. All other authors declared no conflicts of interest.

Word count of abstract: 246 words

Word count of manuscript: 4,426 words

Total number of figures and tables: 5

Supplementary document: 3 Tables and 1 Figure

Dose Transition Pathways: The missing link between complex dose-finding designs and simple decision-making

Christina Yap[1], Lucinda J Billingham[1], Ying Kuen Cheung[2], Charlie Craddock[3], John O'Quigley[4]

(1) Cancer Research UK Clinical Trials Unit, University of Birmingham, Birmingham, UK
(2) Department of Biostatistics, Columbia University, US
(3) Centre for Clinical Haematology, Queen Elizabeth Hospital, Birmingham, UK
(4) Université Paris VI, Paris, France

Correspondence to: Christina Yap, c.yap@bham.ac.uk

Abstract

The ever increasing pace of development of novel therapies mandates efficient methodologies for assessment of their tolerability and activity. Evidence increasingly support the merits of model-based dose-finding designs in identifying the recommended Phase II dose compared to conventional rule-based designs such as the 3+3 but despite this, their use remains limited. Here, we propose a useful tool, Dose Transition Pathways (DTP), which helps overcome several commonly-faced practical and methodological challenges in the implementation of model-based designs. DTP projects in advance the doses recommended by a model-based design for subsequent patients (stay, escalate, de-escalate or stop early), using all the accumulated information. After specifying a model with favourable statistical properties, we utilise the DTP to fine-tune the model to tailor it to the trial's specific requirements that reflect important clinical judgements. In particular, it can help to determine how stringent the stopping rules should be if the investigated therapy is too toxic. Its use to design and implement a modified Continual Reassessment Method is illustrated in an Acute Myeloid Leukaemia trial. DTP removes the fears of model-based designs as unknown, complex systems and can serve as a handbook, guiding decision-making for each dose-update. In the illustrated trial, the seamless, clear transition for each dose-recommendation aided the investigators' understanding of the design and facilitated decision-making to enable finer calibration of a tailored model. We advocate the use of the DTP as an integral procedure in the co-development and successful implementation of practical model-based designs by statisticians and investigators.

INTRODUCTION

The majority of Phase I dose-finding trials have been carried out using conventional rule-based designs, such as the 3+3 design typically seen in oncology trials(1,2). However, in recent years, there have been calls to put the 3+3 Design to rest, particularly in the contemporary era of dose-finding trial settings involving targeted therapy and combination therapy(3-5). There is consequently an emerging interest in more innovative and versatile model-based dose-finding methods such as the Continual Reassessment Method (CRM)(6) and Escalation with Overdose Control(7), that can be extended to cope with more complex clinical settings to meet the demands of the rapidly changing field of novel therapeutics(4). Such approaches aim to utilise all accumulated data to make informed decisions on dose recommendation during the trial as well as the final recommended dose for the next phase. Despite their apparent increased statistical complexity, they have demonstrated the following advantages in comparison to rule-based designs within the statistical literature:

- Superior performance in correctly identifying the right dose (primary trial objective) and hence enable a quicker and more effective progression to later phases of clinical development
- Ability to expose fewer patients to potentially toxic doses and
- Allocation of more patients to desirable dose(s)(8).

Nevertheless, adoption of these advanced methods has been very limited due to the perceived challenges in the implementation of such complex designs(9). Firstly, despite the flexibility that such adaptive designs offer, there exist methodological challenges such as how to choose an appropriate design that takes into consideration the specific requirements of the trial, and is applicable in practice. Secondly, such designs are less familiar than rule-based designs perceived to be “successful” for decades in determining a safe dose. It thus remains a challenge to communicate to investigators (including clinicians and the trials’ management team) how the design works as well as the rationale for using a more resource-intensive and statistically complex design. Thirdly, operational challenges include trialists’ impression that dose-recommendations come from an unknown system (“black-box”), which contrast unfavourably with the transparent simple rules of a rule-based design. There are concerns too about possible delays that might result from the operational team needing to await complex statistical analysis to recommend the next dose. Close collaboration between the statistician and investigators will also be required throughout the dose-finding phase. Such infrastructure may only be present at centres which are able to

invest on specialised statisticians with the relevant expertise (10). This is a major contrast to the simple operational aspect of a rule-based design, which can and has been conducted without a statistician's involvement. Its dose-recommendations are readily anticipated without the need for any kind of calculation.

In this paper, we introduce Dose Transition Pathways (DTP) as a practical tool to aid design, analysis and operation of a trials-specific tailored model that is applicable in practice. We first focus on dose-finding trials whereby the primary objective is to determine the Maximum Tolerated Dose, MTD. This is commonly monitored via a binary variable of occurrence of dose-limiting toxicities (DLT) within a specified assessment period, as defined in the trial protocol. Extension to other model-based designs will be discussed later.
The use of the DTP will be illustrated in the Viola trial - a Phase I trial in Acute Myeloid Leukaemia (AML), where a modified CRM is implemented.

METHODS

The idea of DTP is to project in advance the recommended doses for subsequent cohorts (stay, escalate or de-escalate or stop the trial early), depending on the information accrued thus far.

In a characteristic CRM design, certain parameters need to be considered in the choice of a suitable design. Figure 1 displays the typical specification to be considered for a model-based design such as the CRM(11,12). This can be broken down into the broad categories of clinical parameters, model specification parameters and practical considerations. It highlights that due to the flexibility of such designs, the number of input parameters to consider is in stark contrast to what is necessary for a rule-based design. The latter typically considers only the clinical parameters category, including DLT definition and assessment period, set of doses, fixed cohort size and starts at the lowest dose, and has pre-set dose-escalation rules. On the other hand, one needs to deliberate how the decisions on the numerous parameters for model-based designs (e.g. initial guesses of DLT rates or stopping early) would impact on overall performance and dose-recommendation.

Using the Dose-Transition Pathways as a Design Calibration Tool

Figure 2 displays the practical steps that could be considered in using DTP as a design calibration tool. Several studies have investigated ways to calibrate a CRM design in order to optimize performance(11,13,14). As with all adaptive designs, it is important to conduct statistical simulations to examine the average performance (operating characteristics) of the chosen design under several clinically relevant scenarios (Step 1, Figure 2)(12).

One might argue that even after extensive simulations to select a statistically optimal design under certain conditions, there is a risk it might not be translatable in practice. For instance, during its implementation, the clinical investigators may decide to deviate from what the model recommends, if it recommends doses that are contrary to what they expect. Though use of such designs is to provide guidance for dose-recommendation and is non-binding, repeated deviations from the model would undermine confidence in it. This can be avoided by utilising the DTP as a visual tool to provide better insights – to statisticians as well – than a set of unintuitive statistical equations. This gives us the ability to know when our models are useful and when they might not be appropriate (Step 2, Figure 2).

Here, we introduce the idea of "reverse engineered decision" with an objective to determine how stringent the stopping rules should be to stop a trial early if the lowest dose is too toxic. There are certain pathways whereby the clinical investigators are certain they would recommend stopping if they do observe unexpected levels of DLTs, particularly in the initial cohorts when safety data are limited, e.g. 2/3 or 3/3 DLTs at the lowest dose, or even 1/3 DLT if the acceptable DLT level is low. An example of a Bayesian safety stopping early criterion is displayed below:

p(true DLT rate at lowest dose > target DLT rate + x | current observed data and any relevant prior information) > y

This implies if there is a high chance (> y) that the true DLT rate at the lowest dose is greater than (target DLT rate + x), we would recommend stopping the trial early. The threshold value y is usually calibrated to obtain a design with good operating characteristics (stopping early if all doses are too toxic), but here we will utilise the DTP to fine-tune it to ensure the design also reflects how clinical decisions are likely to be undertaken particularly if unanticipated toxicities arise in the initial cohorts. To aid discussion on when one would

stop early, the DTP can be colour-coded (using a heatmap as employed in the Viola trial described later) to highlight the pathways where there are high risks that that the lowest dose is too toxic.

After obtaining an updated investigator-endorsed DTP of the initial cohorts (Step 3, Figure 2), one then re-run simulations to assess if the model's average performance is still promising as it may be different from the unmodified CRM (Step 1, Figure 2) due to the added modifications. In particular, the impact on the performance of the design under the scenario when all doses of the therapy are too toxic (i.e. the trial should stop early and no dose should be considered suitable as the MTD) should be evaluated. The whole process is completed if the model's operating characteristics are favourable. Otherwise, further calibration might be required before applying it in a trial. The two main processes of simulations to assess operating characteristics and generation of DTP go hand in hand, and the whole procedure may be somewhat iterative before the final design is chosen (Step 4, Figure 2).

We illustrate later how those practical steps in using the DTP were undertaken to fine-tune the CRM model in the Viola trial, with particular focus on Steps 2 and 3.

Using the Dose-Transition Pathways as an Analysis and Operational Tool

The design rules of a 3+3 design dictate that if we observe 0 out of 3 (0/3) patients experiencing a DLT, we will escalate to the next higher dose; if we observe 1/3 DLT, we will include another cohort at the current dose; and if we observe at least 2/3 with DLT, we will de-escalate to the next lower dose or stop the trial if the current dose is the lowest dose. Given the pre-set design rules, the 3+3 design is very straightforward in its operation and no analysis for dose-recommendation is required at the interim stages after each cohort. Often, a statistician may not be required in the running of such designs.

An adaptive model-based design is not as straightforward in its analysis or operational aspect, because the next recommended dose at each interim stage will depend on all accumulated data (i.e. all patients' DLT outcomes at all treated dose levels), as opposed to only recent data. Thus the DTP can be a valuable analytical and operational tool. In a way, it pre-analyses all possible pathways in advance at the planning stage to provide dose-recommendations for subsequent cohorts, and can be displayed in the form of a table or a

flow diagram. With *b* cohorts each including *a* patients giving a total sample size of *ab*, there are $(a+1)^b$ possible pathways. For instance, ten cohorts of three patients would produce up to 1 million possible pathways. The number of unique pathways could be reduced if the trial is stopped early. Displaying the pathways in full is neither practical nor useful. A practical recommendation is to produce a DTP for a specific number of cohorts first. For instance, for a cohort size of three patients, DTP for the first three complete cohorts with up to 64 possible pathways can be produced initially. This would facilitate discussion on the design and its operation with the investigators, which is often the critical stage as data are limited. Once agreement is reached, the DTP based on the adopted tailored CRM model, can form part of the operational procedures for the dose-escalation process. Inclusion of the DTP for the first group of cohorts in the protocol and Statistical Analysis Plan can also facilitate understanding on the workings of such complex designs, overcoming the challenge of navigating through the complexities of novel statistical designs commonly faced by the review committees (12).

The DTP could be updated (using the same tailored CRM model) to project pathways for subsequent (e.g. three) cohorts. In this way, the DTP will only have to be updated in a few stages, rather than after every cohort. Useful discussion on the projected further pathways can be conducted after each update. This may reduce the operational demands of such designs, as well as ensure smoother flow of the trial, avoiding undue delay. It could potentially reduce the need for the statistician to be at hand after each cohort.

Furthermore, the DTP can allow the trial to easily implement the “look ahead” strategy if the next recommended dose by the CRM model is the same regardless of the outcome of the remaining patients in the current cohort. This has an attractive advantage of reducing waiting time between cohorts – a valuable time-saving benefit.

RESULTS

Viola as an illustrative example

Viola (ISRCTN 98163167) is a Phase I trial using a CRM design in AML patients who have relapsed after allogeneic stem cell transplantation. Currently, treatment options for such patients are extremely limited and the great majority die of resistant disease. Small studies showed both azacitidine and lenalidomide possess anti-leukaemic activity when

administered as monotherapy in non-transplant patients whilst azacitidine is active and well tolerated post-transplant(15). Administration of lenolidomide is associated with a significant risk of severe graft-versus-host disease (GvHD)(16). On the basis that azacitidine may moderate the risk of GVHD post-transplant(17), Viola wished to explore the MTD of lenolidomide in combination with azacitidine. Given the lack of effective treatment options currently in the study population, the investigators were happy to accept a higher target DLT probability of 20% (defined as Grade 3/4 or recurrent Grade 2 acute GvHD or most Grade 3 or 4 non-haematological toxicities) in exchange for potentially higher additive efficacy benefit for the combined therapy which has not been tested for this patient population.

The set of seven lenalidomide doses considered for the Viola Trial, ranging from Dose Level -2 to 4, with initial guesses (skeletons) of DLT rates, are displayed in Table 1. Azacitidine was fixed at 75mg/m^2. Patients were recruited in cohorts of 3 with a target sample size of 27. Prior guess of MTD was at Dose Level 1. However, as this was the first time that the two therapies were combined in this patient population, a cautious starting dose level was decided at Dose Level 0. No skipping of untried doses in escalation was allowed whereas skipping of untried doses in de-escalation was permitted. A one-stage, one-parameter empiric Bayesian CRM model was used, with a normal prior of mean 0 for the slope parameter of the dose-toxicity curve. Following the recommended steps as in Figure 2, simulations were first conducted to assess the operating characteristics of the design under several clinically relevant scenarios and a prior variance of 0.75 was selected based on favourable performance (step 1, Figure 2).

Next, we assessed the initial DTP (step 2, Figure 2) for the first three cohorts with all 64 possible pathways (Supplementary Table S1). As mentioned earlier, one could opt to pre-specify all 4^9 possible dose pathways for nine cohorts and store it as an extended spreadsheet; however displaying such a large table upfront may not be practical or useful for discussion or operational purposes. The investigators were in agreement with the model's recommended pathways in terms of escalating, de-escalating or staying at the same dose (Supplementary Table S1 for initial DTP). The one obvious parameter that needed calibration was when to stop the trial early for excessive DLT at the lowest dose. If we observe at least 2 DLTs out of 3 patients at the starting dose (third lowest dose), the estimated DLT rate at the untested lowest dose, Dose -2 is likely to be high. The investigators were unwilling to stop the trial early without first testing the lowest dose, which

is monotherapy of azacitidine. Also, for the first cohort of three patients at the lowest dose, the investigators would only want to stop the trial if there were at least 2 DLTs. Pathways where there were high chances that the lowest dose was too toxic were highlighted using a DTP-heatmap (Supplementary Figure S2). Coupled with initial DTP in Supplementary Table S1, this DTP-heatmap served as a useful tool to discuss with the investigators when one should stop the trial early for excessive toxicity by focusing on those scenarios that were red (high risk that the lowest dose was too toxic) and orange (fairly high risk) based on the first three cohorts. This led to fine-tuning of the threshold value of y as 0.72 to produce pathways which reflect the empirical requirements of the investigators on when they desire the model to recommend stopping early (as in Step 3, Figure 2). This implies if there is a high chance (>72%) that the true DLT rate at the lowest dose is more than 30%, the model will recommend early stopping. The Trial Management Group and Trial Steering Committee (TSC) will be alerted and the latter, with support of any external evidence, will recommend if the trial should be stopped.

The operating characteristics of the design were updated to take into account the fine-tuned safety criteria and the operating characteristics remained favourable (Step 4, Figure 2). The final DTP, based on the tailored CRM design, with 52 unique dose pathways for the first three complete cohorts was discussed and agreed by the clinical investigators as well as the TSC before implementation (Table 2). It was included in both the protocol and the Statistical Analysis Plan. From the DTP in Table 2, if we observe 0/3 DLTs for cohorts 1-3, the model will recommend escalation to the next higher dose after each cohort, as expected. If we observe 1/3 or at least 2/3 DLTs for the first cohort, the model will recommend de-escalation to Dose level -1 and Dose level -2 respectively. An alternative flow diagram of DTP is provided in Figure 3. Some of the pathways lead to the same eventual doses and could be combined if preferred.

An illustration where the look ahead strategy can be applied is if we observe 1/3 DLT in Cohort 1 at Dose 0 and treat Cohort 2 at Dose -1 and observe 2 DLTs in patients 4 and 5 (pathways 25-32, Table 2). Regardless of the DLT outcome of patient 6, the recommended dose for the next cohort is de-escalation to Dose -2. Hence, one can avoid the usual recruitment suspension between cohorts whilst assessing DLT and proceed to recruiting patients in Cohort 3.

Once the DLT outcomes are observed for the first three cohorts, the DTP can be updated to project dose pathways for Cohorts 4-6 and subsequently for Cohorts 7-9, using the same tailored CRM model adopted for Cohorts 1-3. In a hypothetical situation where 0/3 DLT in Cohort 1 at Dose 0, 1/3 DLT in Cohort 2 at Dose 1 and 0/3 DLT in Cohort 3 at Dose 1 are observed, the recommended dose for Cohort 4 is Dose 2 (pathway 5, Table 2). Using this accumulated information, the DTP can be readily updated for Cohorts 4-6, as displayed in the Supplementary Table S3. Assuming 1/3 DLT at Dose 2, 0/3 DLT at Dose 1 and 2/2 DLT at Dose 2 for Cohorts 4 to 6 respectively (pathway 19 in Supplementary Table S3), subsequent DTP for Cohorts 7-9 and the final recommended MTD can be produced, taking into account all the accrued information at all doses (Supplementary Table S4).

As the software code for producing the DTP can be written upfront at the design stage, the update is straightforward and can be easily prepared. In the Viola trial, with the use of the DTP, this allows the user the option to only update twice, after Cohorts 3 and 6. In contrast, using the CRM without the DTP will require eight updates (after every cohort) for dose-assignments. This reduces the operational demands and resources (particularly real-time statistical support) typically required for such complex designs.

DISCUSSION

It is well established that model-based designs are superior to rule-based designs. However, we recognise that implementation of CRM faces several challenges in the design, analysis and operational stages. We thus present DTP as a solution for these challenges.

A potential disadvantage of model-based designs in general, compared to rule-based designs, is their lack of transparency. The key role of the DTP is to bridge this gap by translating the opaque model-based recommendation into numerical decisions that clinicians can comprehend and be able to make inputs to further calibrate the model, ensuring any unexpected behaviour can be resolved prior to implementation. In particular, it helps to determine how stringent the stopping rules should be to stop a trial early if all doses are too toxic. Operationally, this is an attractive tool, as the visualisation of the pre-specified pathways on how the model will recommend doses depending on the accumulated data takes away the fears of such designs being black-boxes. It can serve as a handbook of dose pathways adopted in advance by the investigators, guiding decision-

making for each dose update and reducing the need for constant real-time statistical support similar to the operation of a simple rule-based design.

Furthermore, DTP can be a useful tool to facilitate the work of the relevant safety monitoring board as the dose pathways are clearly displayed in advance, and any concerns could be raised prior to the start of the trial, or after each stage of the updated DTP. If the trial happens as planned with no undue concerns, recruitment of the next cohort can be continued at the model's recommended doses as agreed beforehand. This may reduce the need for safety overview committees to meet regularly after each cohort because of the otherwise opaque dose-recommendation of an adaptive design, speeding up the process to proceed to recruiting the next cohort. Even in unexpected circumstances where there is a deviation in the design (e.g. change in cohort size or incorrect dose given), the DTP can be easily updated for future dose-recommendation.

An alternative simpler model-based approach which combines Bayesian based methods with simple up and down rules similar to the 3+3 algorithm is the modified toxicity probability interval, mTPI (18,19). The toxicity rates are modelled independently with a monotone dose-toxicity curve imposed only at the end of the study. The authors proposed a similar idea as the DTP using a simple Excel spreadsheet to project dose decisions in advance for mTPI at the current dose. Recently, another toxicity probability interval method, Bayesian Optimal Interval Design, BOIN (20) also promoted the use of pre-specification of dose-decision table. Notably, one of the main differences in the dose-assignment decisions between the three approaches is that whilst dose decisions (de-escalate, escalate or stay) under mTPI and BOIN are based on number of DLTs amongst number of evaluable patients at the current tested dose, dose decisions under the CRM are dependent on accumulated data at all doses (not only at the current tested dose) thus borrowing strength from across doses.

Horton et al (2017) (21) compared the CRM with mTPI and BOIN via an extensive simulation study. The paper found that the CRM outperformed the latter two competing methods and its superior performance was more pronounced as the number of dose levels increased. As noted by the authors, CRM is more difficult to implement in practice compared to both mTPI and BOIN (where dose assignments can be displayed in advance) as it is less transparent (21) – without the use of DTP. With DTP, more complex but nonetheless more flexible and superior model-based designs, such as the CRM, can be

made more transparent with the projection of future dose pathways, increasing their potential uptake. Also, by facilitating understanding and interaction with clinicians on the workings of such complex designs, the flexibility of such designs, including incorporation of useful clinical judgements, could be more fully utilised in the development of an eventual tailored model, as demonstrated in Viola.

Although we have focused on the demonstration of DTP in relation to CRM, the concept of DTP is equally applicable and useful to inform the decision-making process and design for other model-based designs such as time-to-event CRM (TITE-CRM) for delayed and cumulative toxicity(22,23), escalation with overdose control(7), phase I/II methods that incorporate activity and toxicity to determine the most desirable or optimal dose(24-26), and methods that consider toxicity severity (as opposed to a binary DLT)(27). The user may have to choose a feasible approach to present projected dose pathways in different settings that will be useful to aid decision making.

For instance, in the case of TITE-CRM(22), where we might have full and partial followed-up DLT observation periods for patients, the projected DTP might differ depending on the amount of information available for each dose update. This will be dependent on patients' accrual times and the time of update. It is hence harder to map out the DTP in advance. However, in such a setting, it will still be useful to produce the DTP for the equivalent CRM where the full assessment period is assumed when designing the trial. This can serve as guidance on whether a waiting period should be imposed if dose-decision based on partial information differs from complete information. For joint evaluation of efficacy and toxicity, the DTP will take into account both outcomes for each patient. Each patient can achieve four possible outcomes: (Toxicity, Efficacy), (Toxicity, No Efficacy), (No Toxicity, Efficacy), (No Toxicity, No Efficacy). If the model is going to be updated after every patient (i.e. cohort size of 1), projected DTP for the first three complete patients will produce up to 4^3 = 64 pathways. On the other hand, if the cohort size is 3, it may only be feasible to project one cohort in advance, giving 20 possible pathways. Projection of two cohorts' joint outcomes will produce 400 (=20^2) pathways which might not be useful. To date, we have utilised the DTP to design several model-based dose-finding trials, using CRM(28,29) and using a model which jointly evaluates efficacy and toxicity, EffTox(30). For computations in practice, availability of R codes for generating DTP tables and figures are available from the authors upon request. Also, the DTP can easily be implemented using existing software

(e.g., the “dfcrm” package in R (31)) to inform all possible outcomes in the following cohort(s).

Calibrating a model is key to the success of using a model-based dose-finding design. Simulation is a common tool to evaluate the statistical properties of a selected design over the long run – this can be viewed as the *Big Picture*. However, it is vital too to focus on the finer, *Small Details*, to ensure that the model operates in ways that adhere to the investigators’ judgement, such as coherence in escalation/de-escalation(11,32), stopping early for excessive toxicity and any trial’s explicit requirements. The utility of this tool is even more important at the initial stage where patient numbers are low, by fine-tuning the model to ensure that the dose decisions be guided (arguably sometimes more) by clinical judgement in conjunction with the statistical model. One can then allow a gradual shift of responsibility to the statistical model as data mature, though clinical input is still important. This can help to ensure that even though there are several potential models with promising statistical properties, the selected design (with its individual performance) is one the clinical investigators find reliable and are comfortable to adopt in practice.

A simple, comprehensive, investigator oriented tool to link between complex models and simple decision-making, fulfils a vital and unmet need for better efficiency that can only be achieved by using such designs in dose-finding trials. By assimilating opinions of clinicians, statisticians and trials operational team to co-develop designs which are applicable in practice and can be easily implemented, DTP can serve as an integral step towards greater confidence and increase uptake of such efficient designs. This will undoubtedly contribute to more rapid and effective progression to later phases of clinical development and ultimately, accelerating patients’ access to potentially life-changing innovations.

Acknowledgements

CY and LJB were funded by grant C22436/A15958 from Cancer Research UK. We acknowledge the support of the Trials Acceleration Programme (funded by Bloodwise) for funding the Viola trial which was used as a case study (13019). Celgene has provided free drug and an unrestricted educational grant for the trial. We would like to thank the editor and referees for their comments which have helped to improve the manuscript.

Contributors

CY initiated and developed the concept, analysed and wrote the manuscript. CY and CC designed the Viola trial. LJB, CC, YKC and JO contributed to the development of the concept, interpretation, drafting and finalising of the manuscript. All authors approved the final draft.

Conflicts of interest

Dr. Yap reports personal fees from Celgene (Honoraria), outside the submitted work; Prof. Craddock reports personal fees from Celgene (Honoraria), personal fees from Celgene (Speaker Panel), grants from Celgene (Research funding), outside the submitted work. All other authors declared no conflicts of interest.

References

1. Storer BE. Design and analysis of phase I clinical trials. Biometrics **1989**:925-37.
2. Rogatko A, Schoeneck D, Jonas W, Tighiouart M, Khuri FR, Porter A. Translation of innovative designs into phase I trials. Journal of Clinical Oncology **2007**;25(31):4982-6.
3. Nie L, Rubin EH, Mehrotra N, Pinheiro J, Fernandes LL, Roy A*, et al.* Rendering the 3+ 3 design to rest: more efficient approaches to oncology dose-finding trials in the era of targeted therapy. AACR; 2016.
4. Harrington JA, Wheeler GM, Sweeting MJ, Mander AP, Jodrell DI. Adaptive designs for dual-agent phase I dose-escalation studies. Nature Reviews Clinical Oncology **2013**;10(5):277-88.
5. Sharma RA, Plummer R, Stock JK, Greenhalgh TA, Ataman O, Kelly S*, et al.* Clinical development of new drug-radiotherapy combinations. Nature reviews Clinical oncology **2016**;13(10):627-42.
6. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. Biometrics **1990**:33-48.
7. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. Statistics in medicine **1998**;17(10):1103-20.
8. Iasonos A, Wilton AS, Riedel ER, Seshan VE, Spriggs DR. A comprehensive comparison of the continual reassessment method to the standard 3+ 3 dose escalation scheme in Phase I dose-finding studies. Clinical Trials **2008**;5(5):465-77.
9. Paoletti X, Ezzalfani M, Le Tourneau C. Statistical controversies in clinical research: requiem for the 3 + 3 design for phase I trials. Ann Oncol **2015**;26(9):1808-12 doi 10.1093/annonc/mdv266.
10. Wong KM, Capasso A, Eckhardt SG. The changing landscape of phase I trials in oncology. Nature Reviews Clinical Oncology **2016**;13(2):106-17.
11. Cheung YK. Dose finding by the continual reassessment method. CRC Press; 2011.
12. Iasonos A, Gönen M, Bosl GJ. Scientific Review of Phase I Protocols With Novel Dose-Escalation Designs: How Much Information Is Needed? Journal of Clinical Oncology **2015**;33(19):2221-5.
13. Lee SM, Cheung YK. Model calibration in the continual reassessment method. Clinical Trials **2009**;6(3):227-38.
14. Wages NA, Conaway MR. Specifications of a continual reassessment method design for phase I trials of combined drugs. Pharmaceutical statistics **2013**;12(4):217-24.
15. Craddock C, Labopin M, Robin M, Finke J, Chevallier P, Yakoub-Agha I*, et al.* Clinical activity of azacitidine in patients who relapse after allogeneic stem cell transplantation for acute myeloid leukemia. Haematologica **2016**;101(7):879-83 doi 10.3324/haematol.2015.140996.
16. Mollgard L, Saft L, Treppendahl MB, Dybedal I, Norgaard JM, Astermark J*, et al.* Clinical effect of increasing doses of lenalidomide in high-risk myelodysplastic syndrome and acute myeloid leukemia with chromosome 5 abnormalities. Haematologica **2011**;96(7):963-71 doi 10.3324/haematol.2010.039669.
17. Craddock C, Jilani N, Siddique S, Yap C, Khan J, Nagra S*, et al.* Tolerability and Clinical Activity of Post-Transplantation Azacitidine in Patients Allografted for Acute Myeloid Leukemia Treated on the RICAZA Trial. Biol Blood Marrow Transplant **2016**;22(2):385-90 doi 10.1016/j.bbmt.2015.09.004.
18. Ji Y, Liu P, Li Y, Bekele BN. A modified toxicity probability interval method for dose-finding trials. Clinical Trials **2010**:1740774510382799.
19. Ji Y, Wang S-J. Modified toxicity probability interval design: a safer and more reliable method than the 3+ 3 design for practical phase I trials. Journal of Clinical Oncology **2013**;31(14):1785-91.
20. Yuan Y, Hess KR, Hilsenbeck SG, Gilbert MR. Bayesian optimal interval design: a simple and well-performing design for phase I oncology trials. Clinical Cancer Research **2016**;22(17):4291-301.
21. Horton BJ, Wages NA, Conaway MR. Performance of toxicity probability interval based designs in contrast to the continual reassessment method. Statistics in Medicine **2017**;36(2):291-300.
22. Cheung YK, Chappell R. Sequential Designs for Phase I Clinical Trials with Late‐Onset Toxicities. Biometrics **2000**;56(4):1177-82.
23. Bekele BN, Thall PF. Dose-finding based on multiple toxicities in a soft tissue sarcoma trial. Journal of the American Statistical Association **2004**;99(465):26-35.
24. Yuan Y, Yin G. Sequential continual reassessment method for two‐dimensional dose finding. Statistics in medicine **2008**;27(27):5664-78.

25. Braun TM. The bivariate continual reassessment method: extending the CRM to phase I trials of two competing outcomes. Controlled clinical trials **2002**;23(3):240-56.
26. Thall PF, Cook JD. Dose‐Finding Based on Efficacy‐Toxicity Trade‐Offs. Biometrics **2004**;60(3):684-93.
27. Lee SM, Cheng B, Cheung YK. Continual reassessment method with multiple toxicity constraints. Biostatistics **2011**;12(2):386-98 doi 10.1093/biostatistics/kxq062.
28. Yap C, Craddock C, Collins G, Khan J, Siddique S, Billingham L. Implementation of adaptive dose-finding designs in two early phase haematological trials: clinical, operational, and methodological challenges. Trials **2013**;14(1):O75.
29. Cole M, Stocken D, Yap C. A pragmatic approach to the design and calibration of a Bayesian CRM dose finding trial. Trials **2015**;16(2):P210.
30. Khan J, Yap C, Clark R, Fenwick N, Marin D. Practical implementation of an adaptive phase I/II design in chronic myeloid leukaemia: evaluating both efficacy and toxicity using the EffTox design. Trials **2013**;14(S1):P20.
31. Cheung Y. dfcrm: Dose-finding by the continual reassessment method. R package version **2013**:02-2.
32. Cheung YK. Coherence principles in dose-finding studies. Biometrika **2005**:863-73.

Table 1: Set of dose levels considered in the Viola trial

Dose Level	Azacitidine dose	Lenalidomide dose	Prior Probability (skeleton) of DLT
-2	75mg/m^2	N/A	0.03
-1	75mg/m^2	2.5 mg	0.07
0 (starting dose)	75mg/m^2	5 mg	0.12
1	75mg/m^2	10 mg	0.2
2	75mg/m^2	15 mg	0.3
3	75mg/m^2	25 mg	0.4
4	75mg/m^2	50 mg	0.6

Table 2 Dose Transition Pathways (DTP) for the first three complete cohorts for the Viola Trial using a modified CRM with cohort size of three. It presents the recommended projected doses for the first three complete cohorts giving in each case the recommended dose for cohort 4 with 52 unique pathways. Starting Dose is Dose 0 (75 mg/m^2 Azacitidine and 5mg Lenalidomide).

Pathway	Cohort 1		Cohort 2		Cohort 3		Cohort 4
	Dose	DLT	Dose	DLT	Dose	DLT	Dose
1	0	0	1	0	2	0	3
2	0	0	1	0	2	1	2
3	0	0	1	0	2	2	1
4	0	0	1	0	2	3	0
5	0	0	1	1	1	0	2
6	0	0	1	1	1	1	0
7	0	0	1	1	1	2	-1
8	0	0	1	1	1	3	-2
9	0	0	1	2	-1	0	0
10	0	0	1	2	-1	1	-1
11	0	0	1	2	-1	2	-2
12	0	0	1	2	-1	3	-2
13	0	0	1	3	-2	0	-1
14	0	0	1	3	-2	1	-2
15	0	0	1	3	-2	2	-2
16	0	0	1	3	-2	3	**STOP**
17	0	1	-1	0	0	0	1
18	0	1	-1	0	0	1	0
19	0	1	-1	0	0	2	-1
20	0	1	-1	0	0	3	-2
21	0	1	-1	1	-1	0	-1
22	0	1	-1	1	-1	1	-2
23	0	1	-1	1	-1	2	-2
24	0	1	-1	1	-1	3	-2
25	0	1	-1	2	-2	0	-2
26	0	1	-1	2	-2	1	-2
27	0	1	-1	2	-2	2	-2
28	0	1	-1	2	-2	3	**STOP**
29	0	1	-1	3	-2	0	-2
30	0	1	-1	3	-2	1	-2
31	0	1	-1	3	-2	2	**STOP**
32	0	1	-1	3	-2	3	**STOP**
33	0	2	-2	0	-2	0	-1
34	0	2	-2	0	-2	1	-2
35	0	2	-2	0	-2	2	-2
36	0	2	-2	0	-2	3	**STOP**
37	0	2	-2	1	-2	0	-2
38	0	2	-2	1	-2	1	-2
39	0	2	-2	1	-2	2	**STOP**
40	0	2	-2	1	-2	3	**STOP**
41	0	2	-2	2	**STOP**	NA	**STOP**
42	0	2	-2	3	**STOP**	NA	**STOP**
43	0	3	-2	0	-2	0	-2
44	0	3	-2	0	-2	1	-2
45	0	3	-2	0	-2	2	-2
46	0	3	-2	0	-2	3	**STOP**
47	0	3	-2	1	-2	0	-2
48	0	3	-2	1	-2	1	-2
49	0	3	-2	1	-2	2	**STOP**
50	0	3	-2	1	-2	3	**STOP**
51	0	3	-2	2	**STOP**	NA	**STOP**
52	0	3	-2	3	**STOP**	NA	**STOP**

Figure 1. Design considerations for Continual Reassessment Method and similar approaches. (Rule-based designs typically consider only the clinical parameters category.)

Figure 2. Practical steps in using DTP to fine-tune model-based designs at the design stage, with the integration of investigators' opinions in the decision-making process, to produce a tailored model.

Figure 3. DTP in the form of a flow diagram for projected dose pathways for Viola. It displays the starting dose as d(0) for Cohort 1 (C1) and the corresponding projected recommended doses for Cohort 2 (C2) and Cohort 3 (C3) depending on the number of DLTs observed, for a cohort size of three. The coloured lines of green, amber and red indicate escalation, staying at the same dose level and de-escalation respectively. The filled red boxes with "STOP" indicate stopping the trial early due to strong evidence that the lowest dose is too toxic.