available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.ejconline.com

Using the continual reassessment method: Lessons Learned from an EORTC phase I dose finding study

Xavier Paoletti[a,*], *Benoît Baron*[b], *Patrick Schöffski*[c], *Pierre Fumoleau*[d], *Denis Lacombe*[b], *Sandrine Marreaud*[b], *Richard Sylvester*[b]

[a]*Inserm, U738, F-75018 Paris, France; Bichat-Cl. Bernard University Hospital, DEBRC, F-75018 Paris, France*
[b]*European Organization for Research and Treatment of Cancer, Data Center, B-1200 Brussels, Belgium*
[c]*Department of General Medical Oncology, Universitair Ziekenhuis Gasthuisberg, B-3000 Leuven, Belgium*
[d]*Department of Medical Oncology, Centre Georges-Francois-Leclerc, F-21079 Dijon, France*

ARTICLE INFO

Article history:
Received 16 November 2005
Received in revised form 22 December 2005
Accepted 3 January 2006
Available online 5 June 2006

Keywords:
Maximum tolerated dose
Dose-toxicity model
Composite endpoint
Dose finding

ABSTRACT

Many clinicians often do not feel comfortable with the Continual Reassessment Method (CRM). This article reviews its implementation, showing the characteristics, advantages and limitations of this method in Phase I studies as an alternative to the classical 'Fibonacci' escalation schema. A two center, dose escalation phase I study of rViscumin was carried out. Thirty-seven patients were included at 14 different dose-levels (10 to 6400 ng/kg). The complete clinical results are presented elsewhere. A 2-step CRM design enables one to speed-up the study and most importantly to obtain an accurate estimate of the maximum tolerated dose (MTD). Different management issues related to a multicenter study are illustrated and we show how the method can go wrong when severe toxicity, or dose limiting toxicity (DLT), is not considered by the clinician as being sufficient to limit dose escalation (here a grade 3 asthenia related to the drug). This would have affected any dose finding methods. We believe that CRM is a good alternative to the standard method from both a statistical and a practical point of view but further methodological research is necessary to address the issues related to the composite nature of the endpoint.

1. Introduction

The objectives of phase I dose escalation studies are to investigate the safety of a compound and to determine the Maximum Tolerated Dose (MTD) while at the same time giving patients the greatest chance of clinical benefit.[1] A good design should take into account the following stringent requirements:[2]

- to not under-treat patients due to the risk of absence of any anti-tumour activity;
- to not over-treat patients due to the risk of severe toxicity;
- to minimize the sample size as the agent under study has not yet been shown to have any anti-tumour activity;
- to gather as much information as possible on the toxicity profile at the MTD;
- to evaluate the pharmacokinetic (PK) profile of the compound.

The Continual Reassessment Method (CRM) is a dose finding method that is increasingly used to conduct phase I studies of cytotoxic agents.[3–6] However to clinicians it may appear

* *Corresponding author:* Present address: Institut National du Cancer, Département Recherche Clinique et Biostatistique, 21 rue Leblanc - Le Ponant B - 75040 Paris Cedex 15. Tel.: +33 (0)1 53 98 54 96; fax: +33 (0)1 53 98 52 49.
E-mail address: xpaoletti@institutcancer.fr (X. Paoletti).
0959-8049/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2006.01.051

sophisticated and difficult to use.[7] Too few methodological publications in the medical literature have highlighted the practical aspects of this method.[8]

Based on the original statistical papers as well as recent developments, we explain the principles and the properties of this method. We illustrate how we put CRM into practice in the study of a new lectin carried out by two institutions of the former New Drug Development Group of the EORTC. rViscumin (Viscum AG) is a new therapeutic strategy in the armamentarium against malignancies which has both immunostimulatory and cytotoxic properties. The drug was administered twice weekly intravenously in patients with solid tumours after failure of standard treatment. The study was approved by the local Human Investigations Committees of both institutions and informed consent was obtained from each subject. A detail report of the clinical trial results can be found elsewhere.[9]

The next section presents the methodological aspects of the study. In a third section, we describe the practical application of this method as well as different management issues encountered during the course of the study. In particular, we underline the limitations of the method when sporadic Dose Limiting Toxicity (DLT) occurs at a dose-level which is much lower than the actual MTD.

2. Patients and Methods

2.1. *Endpoint and objectives*

The main objective of this study was to identify the MTD based on the occurrence of DLTs during the first cycle of treatment. DLTs consisted of any haematological grade IV or non-haematological grade III–IV adverse event according to the Common Toxicity Criteria Version 2, with the exclusion of nausea, vomiting or fever which could be rapidly controlled.

The proportion of patients who suffer at least one DLT related to treatment during the first cycle was expected to increase with the dose. The MTD is the dose closest to a pre-specified "acceptable" risk of DLTs, here 20%. For illustrative purposes, consider Fig. 1 which represents a hypothetical increasing dose-toxicity relationship (solid line). The dashed line indicates the 20% target. The dose associated with the MTD is then d_4. If a different level of acceptable risk for a DLT is chosen, for instance 40% (dotted line), the MTD would then be dose-level 5. The target risk depends on what level of toxicity is considered to be acceptable in a given patient population and type of treatment.

2.2. *Dose*

A starting dose of 10 ng/kg was taken as 1/100 of the MTD in dogs. Dose increments were fixed in advance, but the number of doses to be investigated was not specified. The dose was doubled as long as only grade 0 or 1 toxicity was encountered or until a dose of 1600 ng was reached. In the absence of toxicity at a dose of 1600 ng, the subsequent dose increments were +800 ng. This gave the following dose sequence: 10, 20, 40, 100, 200, 400, 800, +800... ng.

2.3. *Statistical design*

The experimental design for this study was a simple application of the likelihood version of the CRM. Therefore no prior distribution was required for the parameter "a" defined below. The design was split in two consecutive steps[10]: an escalation step and a model guided step. Both were clearly defined in the protocol. The escalation step enables to quickly clear low-dose-levels which are likely to be safe, but with poor anti-tumour activity, without over-shooting the MTD. The study started at the lowest dose. Since very low toxicity was expected, a cohort of 1 patient was treated at level 1. As long as toxic side-effects were grade 0 or 1, dose-levels were escalated after each inclusion. In the event of moderate drug-related toxicity (which was not a DLT), the cohort size was increased from 1 to 3 patients. If none of the 3 patients experienced a DLT, it was possible to proceed to the next dose-level. As soon as the first DLT was observed, the escalation step was interrupted, and the model guided step kicked in.

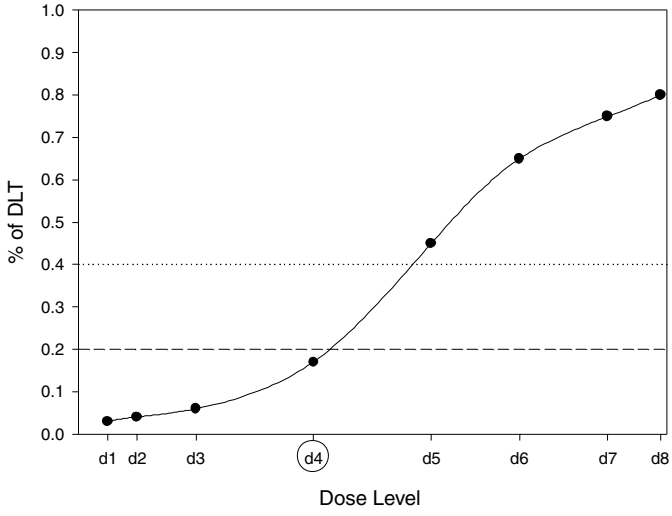

Fig. 1 – Hypothetical dose-toxicity relationship.

The dose-level to be administered to the next patient was then reassessed after each new inclusion given all previous data (cohort size of 1). Continual reassessment assures the best operating characteristics and also gives the patient the greatest chance to be administered at the MTD.

The general principle is to fit a dose-toxicity model to all data after the evaluation of toxicity in each new patient. This provides a continual reassessment of the estimate of the probability of toxicity at each dose. The dose-level for the next patient is then the dose associated with an estimated risk of DLT which is the closest to the target, i.e. 20%. This dose is called the running MTD estimate.

A simple one-parameter model was chosen to link the risk of a DLT to the dose: $P(\text{DLT at } d_i) = \alpha_i^a$, where α_i is a recoding of the dose-levels d_i fixed in the protocol and 'a' the parameter to be estimated. Using codes makes it easier to set-up the model and to obtain desirable operating characteristics (e.g. no skipping of doses, asymptotical convergence).[11] The doses were coded as proposed by O'Quigley and colleagues.[10] They were only specified after the occurrence of the first DLT, since the model is unnecessary before that (see Table 1). There are no absolute rules to determine these codes and this is still an area of research. Simulations are an efficient tool in assessing the properties of the chosen coding.

One of the main characteristics of this method is that as information on more patients become available, the more accurate the MTD estimate and the more precise the safety assessment. In the current study, we implemented the stopping rule proposed by O'Quigley and Reiner,[12] which states that at any point of the study, we estimate the probability that the next 5 patients will be allocated the same dose-level, based on the current estimate of the probability of toxicity at the running MTD. The study stops when this exceeds 90%. The current MTD is then the dose-level recommended for Phase II trials.

2.4. Ethical approval

The trial was approved by the local Human Investigations Committee of the Medizinische Hochschule Hannover and of the Centre Rene Gauducheau in Nantes. The EORTC insurance program covered all patients entered in this EORTC study.

3. Results

3.1. Escalation step

The escalation step from patient 1 to patient 11 is summarized in Fig. 2. The y-axis displays the dose-level attributed to the patient whose number is indicated on the x-axis.

As no moderate toxicity was observed in the first 10 patients, the dose escalation step continued until the first DLT, grade 3 fatigue, was observed in patient 11 at dose d_{11}, at which time the model guided step kicked in. The result was an astonishing dose escalation from 10 ng/kg up to 4000 ng/kg. Obviously, the starting dose was not appropriate and one can imagine what would have happened with a slower dose escalation scheme.

3.2. Model guided step: Patients 12–26

After the first DLT in patient 11, the model was fit as shown in Fig. 3 (as the parameter estimated $\hat{a} \approx 1$, the probability of a DLT seemed linear). This gave the estimated probabilities of toxicity provided in Table 2.

Table 1 – Doses (d_i) and associated codes (α_i)

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}
Dose (ng/kg)	10	20	40	100	200	400	800	1600	2400	3200	4000	4800	5600	6400
α_i	0.0035	0.005	0.009	0.015	0.024	0.035	0.05	0.07	0.11	0.20	0.33	0.48	0.62	0.74

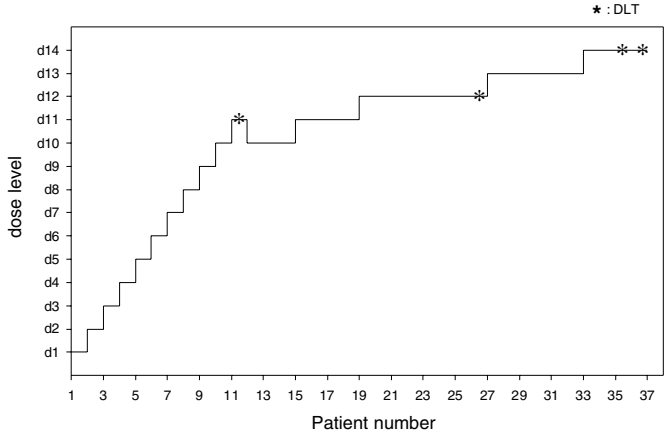

Fig. 2 – Trial history.

The probability estimates of a DLT at d_{10} and d_{11} were respectively 18% and 30%. Since the MTD is defined as the dose associated with 20% of DLTs, d_{10} appeared to be the best candidate after 11 patients and was recommended for patient 12. It is important to highlight that the estimates are rather imprecise due to the small number of patients that have been treated. In particular estimates at dose-level 12 and higher only rely on model extrapolation. Patient 12 was treated at dose d_{10}, 3200 ng/kg, which was well tolerated. This new data was fed into the model and estimates of the probability of a DLT were reassessed, with values 0.15 and 0.27 at levels d_{10} and d_{11} respectively. Patient 13 was treated at the dose indicated by the CRM, d_{10}, and supported the treatment. The recommended dose for patient 14 was d_{11}. However to obtain valuable PK data, it was decided to also treat patient 14 at d_{10}. Patient 15 was then treated at the recommended dose d_{11}. None of these patients had a serious adverse event. The DLT probabilities at d_{10} and d_{11} were then estimated to be 11% and 22%. Patient 16 was administered d_{11} as were patients 17 and 18 based on the model's recommendations. The different plots in Fig. 4 give a good insight into the evolution of the estimated risk of severe toxicity. Panel A shows the effect of non-DLTs. Panel B shows the effect of a DLT later on in the study. The DLT of patient 26 points toward a lower dose, but dose-level 12 is still recommended as being the closest to the target.

In the absence of further DLTs, patients 19 through 26 were treated at dose d_{12} due to little change in the dose limiting toxicity estimations. An important feature of the method appeared when patient 26 was treated at d_{12} and suffered from grade 3 transaminitis, a DLT. After this DLT, the estimated probabilities of a DLT around the running MTD (d_{12}) were: 0.00, 0.01, 0.03, 0.08, 0.19 and 0.33 at d_8 to d_{13} respectively. However, after 8 subjects were treated at dose d_{12} with one DLT, what were the CRM recommendations? To get an overview of future dose allocations and plan future patient entry, an exploratory decision tree was used (see Fig. 5). The vertical scale along the tree diagram is the patient number, the framed squares indicate the dose-levels, and the branches the 2 possible responses (DLT on the left and no DLT on the right) for each patient. This builds up a binary tree, for instance, patient 26 is allocated to dose 12. If (s)he has a DLT, the left branch indicates that patient 27 would be recommended dose-level 12. After another DLT, dose 11 would be recommended for patient 28. Conversely, had patient 26 tolerated the treatment, patient 27 would have been allocated level 12 and so on.

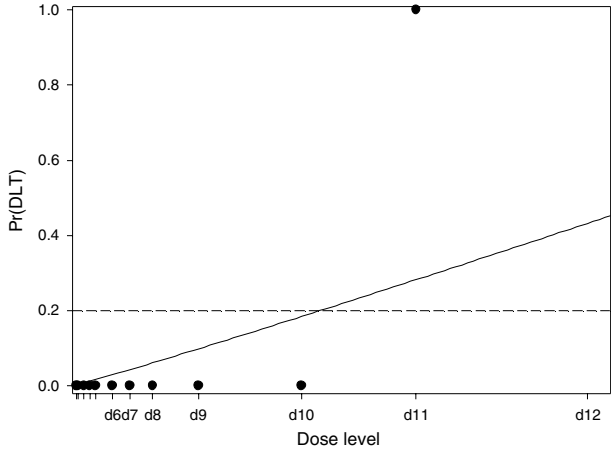

Fig. 3 – Empirical frequencies of DLT and model fit after patient 11. The horizontal line represents the target DLT level (20%), the plain line the fitted power model and the dots correspond to the observed proportion of DLTs.

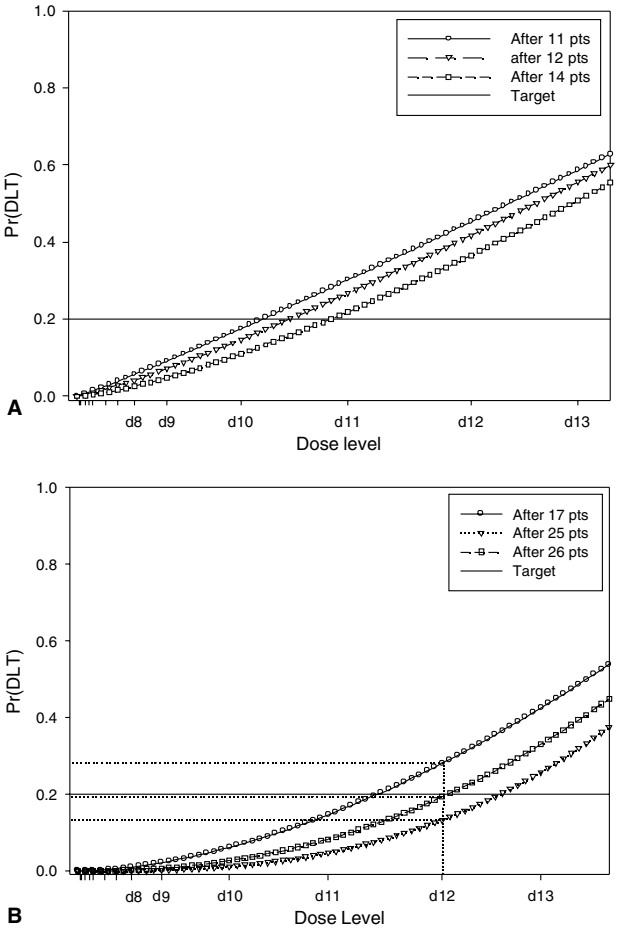

Fig. 4 – (A) Estimated dose-DLT curves after pts 11, 12 and 14. (B) Estimated dose-DLT curves after pts 17, 25 and 26 (DLT).

Table 2 – Probability of a DLT after 11 patients according to dose

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	d_{12}
Dose (ng/kg)	10	20	40	100	200	400	800	1600	2400	3200	4000	4800
P (DLT)	0.00	0.00	0.01	0.01	0.01	0.03	0.04	0.06	0.09	0.18	0.30	0.45

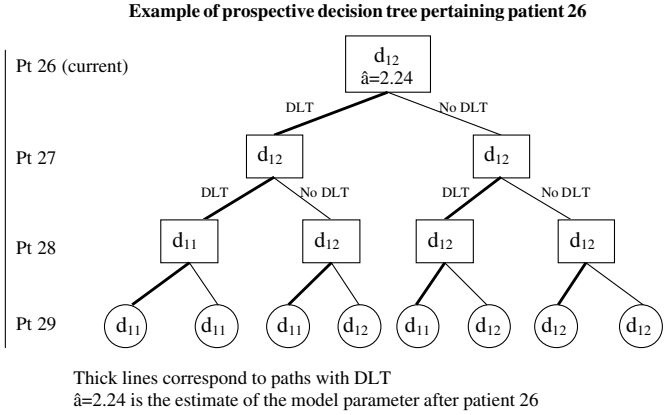

Fig. 5 – Example of prospective decision tree pertaining patient 26.

Table 3 – Final probability of a DLT according to dose (after 37 patients)

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}
Dose	10	20	40	100	200	400	800	1600	2400	3200	4000	4800	5600	6400
P (DLT)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.06	0.16	0.31

Grey column corresponds to the final recommended dose.

After 2 DLTs at d_{11} and d_{12}, following CRM recommendations, 6 additional patients without DLTs would be necessary to escalate to d_{13}. Yet, the first DLT at dose d_{11} was asthenia grade 3, that promptly resolved since patient went back to work in the afternoon following administration, and the second one at dose d_{12} was manageable. Should these toxicities really limit the dose escalation? Actually, for decision making, when coded as DLTs, they have the same value as, for instance, febrile neutropenia. This leads the model to probably over-estimate the global toxicity of the compound under study.

3.3. *Model guided step: Patients 27–37*

Considering the type of adverse reactions observed, the steering committee decided not to follow model's recommendations and to escalate to dose d_{13} for patient 27. It was also decided to reconsider the first DLT and to recode it as a non-DLT. Updated probability estimates were 0.00, 0.01, 0.03, 0.11 and 0.22 at dose-levels d_9 to d_{13} respectively. Patients 27 to 32 were then treated at d_{13}, the dose closest to the target. Starting from patient 33, d_{14} was recommended.

Patients 35 and 37 both suffered from grade 3 transaminitis at a dose of 6400 ng/kg (d_{14}). This led to the final estimates given in Table 3. The final recommended dose was level 13 (5600 ng/kg) at which the estimated probability of DLT was 0.16.

3.4. *Stopping the study*

After patient 37, the exploratory tree analysis showed that the estimated probability to stay at the same dose-level, d_{13}, for the next 5 patients was approximately 90%. The 95% confidence interval of the estimate of the probability of toxicity at d_{13} was [0.07, 0.37] and the type of toxicity was manageable. It was therefore advised to stop the study.

4. Discussion

After 37 patients the study came to a halt and a dose of 5600 ng/kg (560 times the initial dose) was recommended for further investigation. The overall duration of the accrual was 31 months. After an astonishing dose escalation, an isolated DLT slowed down the process and gave the feeling that the method was too conservative. After a second DLT at dose d_{12} (patient 26), investigators chose to reverse the assignment of a DLT to patient 11's grade 3 asthenia. This trial is a good example of the advantages and potential drawbacks of the method.

The two main strengths of this method are its good statistical properties and its flexibility. Its overall performance has been extensively assessed by simulating a large number of possible studies. Several authors have published that CRM gives a greater chance to identify the right dose, in comparison to the standard design.[13–15] They have also found that CRM is better at avoiding treating patients at too low dose-levels and gives patients a greater chance to be treated around the MTD. The risk of over-treatment is higher than with the conservative standard method but due to: i) the integration of moderate toxicity in the decision rule during the escalation stage to slow down dose escalation; and ii) the reassessment after each patient, the risk of treating several consecutive patients at too high a dose-level is limited.

The CRM is quite flexible. The escalation stage fulfilled its role. The initial dose-levels were rapidly cleared without severe toxicity. Using the notion of moderate toxicity to adapt

the escalation speed, together with a dynamic cohort size and dynamic dose increments, probably results in a good trade-off between a fast and a safe design.[13] This is also a straightforward way to integrate toxicity other than DLTs in the decision rule and thus to enrich the design. Moreover, contrary to the usual Fibonacci method where the targeted percentile is invariant around 25%,[16,17] CRM allows one to define MTD differently according to how aggressive the treatment should be. Even though 20% or 25% are the most common probabilities, Thall[11] report a GVHD study at the MD Anderson Cancer Center where a target of 40% was selected considering the important benefit that might be obtained with higher doses.

Another illustration of the versatility of this method relies on the possibility to not strictly follow the method's recommendations but to still rigorously incorporate all the data in the decision rule. This was the case after patient 14 when an extra patient was treated at the same dose in order to analyze PK data. Inferences are based on all available information.

Inferences were drawn by maximizing the likelihood function. An alternative estimation technique, the Bayesian one,[3] requires one to specify a prior distribution for the probability of a DLT. It is often chosen to be non-informative, that is to be easily overridden by the data. Such a specification is difficult when prior information on the range of the doses under study is poorly known. Actually, a non-informative prior over a given range may be strongly informative if the real MTD is far above. In addition, in the absence of DLTs, the dose escalation relies on the prior and *ad hoc* rules which make the design complicated and less flexible. The two-stage CRM using the likelihood technique avoids this pitfall. This flexibility is not only an advantage but also a condition for its good statistical performance.

In particular, if we were to use a fast escalation stage with one patient per dose-level followed by a 3 + 3 design after the first DLT, the risk of having 3 patients treated at too high a dose-level would strongly increase. Storer[16] proposed such a design but he included the possibility to go up, down, and up again if necessary, bringing extra flexibility. He showed this design to be rather efficient. Similarly, Simon and colleagues[18] have shown that the "accelerated titration" design may result in substantial savings in terms of the number of patients treated at low doses at the expense of a slightly higher rate of DLTs. Simulations of this design indicate that the trial would probably have stopped 21 to 24 patients after the first DLT was observed, depending on the toxicity. A further advantage of using a model is the possibility to include covariates. They can be used to explore stratification or the incorporation of PK data[19] which are central in phase I studies. The last aspect of this flexibility is the stopping rule. We definitely think that the usual rule to close the study as soon as 6 patients have been treated at the final dose may often lead to erroneous conclusions. On the other hand, we do not think that the answer to this question should only rely on statistical inference; unlike phase II or phase III trials, a strict rule is probably not the most appropriate. Statisticians can provide investigators with useful information about the accuracy of the MTD estimate, but the final decision should be taken by the steering committee. Statistical tools include the confidence interval of the probability of the DLT estimate[20] or the stopping rule used here.

The overall duration of a trial is often a concern. The CRM is efficient in speeding up the trial during the first stage. However if we focus on the second step of the trial (after first DLT), the CRM and the standard design would have led to a similar duration of the trial. Supposing no delay in patient accrual, 12 months would have been necessary to complete the trial with the standard approach and 14 months with the CRM. The tree enables one to optimize the accrual risk and to anticipate difficulties in escalating the levels. It prospectively provides the investigators with an overview of the future dose recommendations so that they can inform patients in due time. For instance, patient 17 was to be given the same dose whatever the toxicity of patient 16; thus it was possible to accrue both simultaneously. This situation happens more frequently as we go further into the study.

Nevertheless, flexibility reached its limits after the occurrence of the first two DLTs. The difficulties to escalate the dose-levels after the first and second DLTs deserve attention. Several issues should be considered. The first one is the definition of a DLT. Obviously, transient grade 3 asthenia should not have the same weight in the decision process as life threatening toxicities. An important modification of the phase I methodology is necessary in order to take into account differences between the types of DLTs. This is a track followed by Bekele and Thall.[21] The second one is related to the method itself. The strength of the CRM is to concentrate most of the doses around one level where valuable information is collected and valid estimations are possible. However, when observations are spread over different levels, similar to this study due to the two isolated DLTs two levels below the MTD, the estimates become less accurate. It should be noted that standard design would have been less affected since only the last observations are used for the allocation rule. The first solution is to use a richer model together with a Bayesian framework as proposed by Whitehead (1998)[22]; this allows for different patient allocation designs. Another solution could be to weight the DLTs and to modify these weights within the course of the study when more information on the suspicious DLTs becomes available. To some extent we followed this procedure for patient 11. His/her DLT had weight 1 until we recoded it as 0, i.e. a null weight. Without this transformation, the CRM would have led to a lower MTD. This issue concerns all dose finding methods. However model-based methods are probably more appropriate to handle it since they provide a richer framework to reassess responses, to incorporate extra information, and to refine modeling. The relevance of this choice is now confirmed by a follow-up study[23] who observed the same DLTs as in our CRM-driven trial.

This study is a good example of the possibilities and the limitations of CRM method. The classic Fibonacci would have made the study endless. Not only did the CRM enable us to clear 10 dose-levels in a minimum amount of time but it also provides a more precise estimate of the MTD at the end of the trial. This is essential when considering that the MTD is rarely modified in phase II or phase III studies. The scientific objective should not only aim at speeding up the trial, but above all at accurately identifying the optimal dose, which is not

always the case.[24] However, further research needs to be done to refine DLT modeling, to integrate information other than toxicity and to obtain more accurate toxicity probability estimates.

Conflict of interest statement

None declared.

Acknowledgement

We thank VISCUM AG for sponsoring the clinical study of rViscumin. This publication was supported by grant number 5U10 CA11488-35 from the National Cancer Institute (Bethesda, Maryland, USA). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute. This study was presented as a poster at the 2002 ASCO meeting and at the 2002 EORTC-NCI-AACR Symposium.The medical results have been published in the Annals of Oncology in 2004 by Schöffski and colleagues.[6]

REFERENCES

1. Roberts TG, Goulart BH, Squitieri L, et al. Trends in the Risks and Benefits to Patients with Cancer Participating in Phase 1 Clinical Trials. *JAMA* 2004;**292**(17):2130–40.
2. Eisenhauer EA, O'Dwyer PJ, Christian M, et al. Phase I clinical trials design in cancer drug development. *J Clin Oncol* 2000;**18**:684–92.
3. O'Quigley J, Pepe M, Fisher L. Continual Reassessment Method: a practical method for phase I trials. *Biometrics* 1990;**46**:33–48.
4. Geoerger B, Vassal G, Doz F, et al. Dose finding and O6-alkylguanine-DNA alkyltransferase study of cisplatin combined with temozolomide in paediatric solid malignancies. *Br J Cancer.* 2005;**93**(5):529–37.
5. Khuri FR, Glisson BS, Kim ES, et al. Phase I study of the farnesyltransferase inhibitor lonafarnib with paclitaxel in solid tumours. *Clin Cancer Res.* 2004;**10**(9):2968–76.
6. Eder Jr JP, Garcia-Carbonero R, Clark JW, et al. A phase I trial of daily oral 4′- N -benzoyl-staurosporine in combination with protracted continuous infusion 5-fluorouracil in patients with advanced solid malignancies. *Invest New Drugs* 2004;**22**(2):139–50.
7. Goodman SN, Zahurak ML, Piantadosi S. Some Practical Improvements in the Continual Reassessment Method for Phase I Studies. *Stat Med* 1995;**14**(11):1149–61.
8. Dent SF, Eisenhauer EA. Phase I trial design: are new methodologies being put into practice? *Ann Oncol* 1996;**7**:561–6.
9. Schöffski P, Riggert S, Fumoleau P, et al. Phase I trial of intravenous aviscumine (rViscumin) in patients with solid tumours: a study of the European Organization for Research and Treatment of Cancer New Drug Development Group. *Ann Oncol* 2004;**15**(12):1816–24.
10. O'Quigley J, Shen LZ. Continual Reassessment Method: A likelihood approach. *Biometrics* 1996;**52**:395–405.
11. Thall PF, Lee JJ, Tseng CH, et al. Accrual strategies for phase I trials with delayed patient outcome. *Stat Med* 1999;**18**(10):1155–69.
12. O'Quigley J, Reiner E. A stopping rule for the Continual Reassessment Method. *Biometrika* 1998;**85**(3):741–8.
13. Paoletti X, O'Quigley J, Maccario J. *Design* Efficiency in Dose Finding Studies. *Computational Statistics and Data Analysis* 2004;**45**(2):197–214.
14. Ahn C. An evaluation of phase I cancer clinical trial designs. *Stat Med* 1998;**17**:1537–49.
15. Storer B. Choosing a Phase I Design in Handbook of Statistics in Clinical Oncology. In: John Crowley editor, *Handbook of Statistics in Clinical Oncology*. NY, Crowley, 2001, 73–91.
16. Storer BE. Design and analysis of phase I clinical trials. *Biometrics* 1989;**45**(3):925–37.
17. Durham SD, Flournoy N, Rosenberger WF. A random Walk rule for phase I clinical trials. *Biometrics* 1997;**53**:745–60.
18. Simon R, Freidlin B, Rubinstein L, Arbuck SG, Collins J, Christian MC. Accelerated titration designs for phase I clinical trials in oncology. *J Natl Cancer Inst* 1997;**89**(15):1138–47.
19. Piantadosi S, Liu G. Improved designs for dose escalation studies using Pharmacokinetics Measurements. *Stat Med* 1995;**15**:1605–18.
20. Zohar S, Chevret S. The continual reassessment method: comparison of Bayesian stopping rules for dose-ranging studies. *Stat Med.* 2001;**20**(19):2827–43.
21. Bekele BN, Thall PF. Dose Finding Based on Multiple Toxicities in a Soft Tissue Sarcoma Trial. *JASA* 2004;**89**(465):26–35.
22. Whitehead J. A unified theory for sequential clinical trials. *Stat Med.* 1999;**18**(17–18):2271–86.
23. Schoffski P, Breidenbach I, Krauter J, et al. Weekly 24-h infusion of aviscumine (rViscumin): a phase I study in patients with solid tumours. *Eur J Cancer* 2005;**41**(10):1431–8.
24. Christian MC, Korn EL. The limited precision of phase I trials. *J Natl Cancer Inst* 1994;**16**(86):1662–3.